ALGOM: Improvements for GOMOS O2 and H2O profiles retrieval
Technical Note

Ref. : A1299-TN-001-ACR
Version : 2.4
Date : 24/01/2017
Contract ref. : 0166L68540
Signatures

<table>
<thead>
<tr>
<th>Prepared by</th>
<th>Name</th>
<th>Company</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Laurent Blanot</td>
<td>ACRI-ST</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jean-Loup Bertaux</td>
<td>LATMOS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alain Hauchecorne</td>
<td>LATMOS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Francis Dalaudier</td>
<td>LATMOS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Véronique Bruniquel</td>
<td>ACRI-ST</td>
<td></td>
</tr>
</tbody>
</table>

Distribution list

<table>
<thead>
<tr>
<th>Company</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACRI-ST</td>
<td>Laurent Blanot, Véronique Bruniquel</td>
</tr>
<tr>
<td>LATMOS</td>
<td>Jean-Loup Bertaux, Alain Hauchecorne, Francis Dalaudier</td>
</tr>
<tr>
<td>ESA</td>
<td>Angelika Dehn</td>
</tr>
<tr>
<td>FMI</td>
<td>Janne Hakkarainen</td>
</tr>
</tbody>
</table>

Changes Log

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4</td>
<td>23/01/2017</td>
<td>Added the results on SPB ozone correction impact. + Minor changes</td>
</tr>
<tr>
<td>2.3</td>
<td>04/01/2017</td>
<td>Changes according to ESA/SERCO reviews.</td>
</tr>
<tr>
<td>2.2</td>
<td>13/05/2016</td>
<td>Change in the text “+0.1nm” into “+0.015nm (=1/3 pixels)” in limitation – wavelength assignment section. This concern the wl sensitivity analysis test. Other minor changes</td>
</tr>
<tr>
<td>2.1</td>
<td>26/04/2016</td>
<td>Minor changes</td>
</tr>
<tr>
<td>2.0</td>
<td>09/02/2016</td>
<td>Large changes in the document; New results. Detailed results chapter moved in Appendix.</td>
</tr>
<tr>
<td>1.0</td>
<td>04/09/2015</td>
<td>First Version</td>
</tr>
</tbody>
</table>
Table of Content

1 INTRODUCTION .................................................................................................................. 15
  1.1 SCOPE OF THE DOCUMENT .................................................................................. 15
  1.2 STRUCTURE OF THE DOCUMENT ...................................................................... 15
  1.3 APPLICABLE AND REFERENCE DOCUMENTS ...................................................... 15
    1.3.1 Applicable Document ................................................................................ 15
    1.3.2 Reference Document ............................................................................... 15
  1.4 ACRONYMS .............................................................................................................. 16

2 CURRENT ALGORITHM ....................................................................................................... 17
  2.1 DESCRIPTION ......................................................................................................... 17
  2.2 LIMITATIONS ......................................................................................................... 18
    2.2.1 Wavelength assignment .............................................................................. 19
    2.2.2 Spectral Point Spread Function (PSF) ....................................................... 20
    2.2.3 Spectroscopic database ............................................................................. 23

3 NEW ALGORITHM .............................................................................................................. 25
  3.1 DESCRIPTION ......................................................................................................... 25
    3.1.1 Inversion scheme ......................................................................................... 25
    3.1.2 Forward model ......................................................................................... 26
    3.1.3 Limitations ................................................................................................. 27

4 ALGOM DATASET GENERATION ...................................................................................... 30
  4.1 FORMAT AND CONTENT OF THE PRODUCTS ...................................................... 30
  4.2 PROCESSING CONFIGURATION ............................................................................. 32

5 PROJECT SYNTHESIS ...................................................................................................... 33
  5.1 ALGORITHMS DIFFERENCES ........................................................................... 33
  5.2 IMPROVEMENT OF THE QUALITY OF THE PROFILES ........................................ 34
  5.3 ALGOM MISUNDERSTOOD/UNEXPLAINED RESULTS ......................................... 37
  5.4 LEVEL-1B PROBLEMS ......................................................................................... 40

6 FURTHER IMPROVEMENTS .............................................................................................. 42

7 GENERAL CONCLUSIONS ................................................................................................. 43

8 APPENDIX - DETAILED RESULTS ................................................................................... 44
  8.1 O2 ......................................................................................................................... 44
    8.1.1 First results – comparison with GOPR ....................................................... 44
    8.1.2 Dependence with Star brightness in SPB1 ............................................... 54
    8.1.3 Fixing PSF width in LM ............................................................................ 58
    8.1.4 Fixing wavelength assignment in LM ....................................................... 60
    8.1.5 Dependence with HITRAN version ........................................................... 63
    8.1.6 Dependence with CIA ............................................................................... 67
    8.1.7 Dependence with climatological model ................................................... 70
    8.1.8 Build model transmission LUT from Gomos External profile ................. 71

© 2016 ACRI-ST
8.1.9 Understanding the -20 % Negative Bias wrt IPF/GOPR ........................................ 72
8.1.10 Improving the error bars ......................................................................................... 74
8.1.11 False star tracking .................................................................................................. 76
8.1.12 Dark Charge correction .......................................................................................... 78
8.1.13 Use SPA2 O₂B absorption band .............................................................................. 82

8.2 H₂O .............................................................................................................................. 91
8.2.1 First results – comparison with GOPR..................................................................... 91
8.2.2 Fixing PSF width in Levenberg Marquardt. ............................................................... 99
8.2.3 Fixing wavelength assignment in LM ...................................................................... 102
8.2.4 Using retrieved SPB1 wavelength assignment for SPB2 inversion ......................... 103
8.2.5 Dependence with HITRAN version ......................................................................... 106
8.2.6 Dependence with climatological model ................................................................. 109
8.2.7 Problem with iPRNU correction ............................................................................ 110
8.2.8 Comparison ALGOM vs GOPR for the ALGOM-delivered H₂O dataset ................. 111
List of Figures

Figure 1: \( \text{O}_2 \) measured (red) and model (black) transmission spectrum in SPB1. The tangent point altitude of the measured transmission is 14.5 km. The baseline of the measured transmission is lower than 1 because of aerosol and Rayleigh scattering effects. Those effects are not considered in the represented model transmission spectrum. .................................................................17

Figure 2: Scatter plot of \( \log(T_{\text{mes}}) \) vs. \( \log(T_{\text{mod}}) \). The black line represents a linear fit. In the retrieval algorithm, several model transmissions spectra at different altitudes are tested. The altitude model which gives a slope = 1 is considered to be the good one, providing the retrieved \( \text{O}_2 \) line density. ...18

Figure 3: Wavelength assignment shift as a function of orbit number for spectrometer B1 computed from S0002 data by cross-correlation of a low altitude measured atmospheric \( \text{O}_2 \) transmission spectrum vs. a reference spectrum. Measurement spectrum has been taken from the residual extinction product (GOM\_EXT). In blue, the reference spectrum used is a measurement spectrum near orbit R03000. In green, the reference spectrum used is the model spectrum of the current RES\_EXT product but associated with the wavelength assignment near orbit R03000. Both methods give nearly similar results. The blue data has been fitted by a 2\(^{\text{nd}}\) order polynomial (red curve), which provides the ‘imposed’ wavelength assignment shift model used by IPF 6.01 (and GOPR 7) for the processing of all GOMOS occultations (applied to all stars and all seasons). ........................................ 19

Figure 4: Same as Figure 3 but for star S0001. We see strong seasonal variations of amplitude up to 0.05nm. ........................................................................................................................................... 20

Figure 5: Relative difference of \( \text{O}_2 \) tangent line density profiles between: 1. wavelength assignment shift=0 nm (nominal value). 2. Wavelength assignment shift forced to +0.015 nm (= 1/3 pixel). Y-axis is altitude in km. The \( \text{O}_2 \) relative difference reaches about 10% between 40 and 60 km. This study confirms the importance of having a precise wavelength assignment determination for the retrieval of \( \text{O}_2 \) profiles. For this, the seasonal variation of the wavelength assignment shift must be taken into account. ...................................................................................................................20

Figure 6: 3 coloured curves: spectral PSFs measured on ground in thermal vacuum conditions in 1998 for 3 different wavelengths covered by spectrometer B1. Black curve: The Spectral PSF stored in the IPC product (Gaussian). The measured PSFs at 766.5 nm and 769.9 nm have a very similar width whereas the PSF at 756 nm is larger. We note that the IPC PSF is larger than the measured PSFs at 765.5 nm and 769 nm but smaller than the 756.0 nm. There is no \( \text{O}_2 \) absorption lines in SPB1 for wavelengths below 759 nm (see Figure 1), thus a width of the PSF closer to the 766.5 and 769.9 nm values would have been more appropriate to set in the IPC product. This is what has been shown in [RD2]................................................................................................................................. 22

Figure 7: Same as Figure 6 for spectrometer B2. The IPC PSF is the same as for SPB1. It is close to the measured PSF widths. ........................................................................................................................................... 22

Figure 8: From top to bottom: model transmission in SPB1 spectral range (\( \text{O}_2 \) retrieval) for tangent point altitudes of 15, 10, 5 and 1 km without (black) or with CIA (red). Resolving power: 6000, US76. The influence of CIA is almost null at 15 km and above. Source: [RD-3] .................................................. 24

Figure 9: \( \text{O}_2 \) (left) and \( \text{H}_2\text{O} \) (right) slant path density profiles for IPF/GOPR (blue) and ALGOM (red) for all occultations of star S0001. In this plot we see 3 improvements brought by ALGOM: 1. Red circled: GOPR/IPF does not retrieve \( \text{O}_2 \) and \( \text{H}_2\text{O} \) at low altitude while ALGOM does. 2. Green circled: GOPR/IPF has a lot of outliers at low altitudes while ALGOM has not. 3. Black circled: GOPR/IPF has a threshold for lower values while ALGOM has not. ................................................................................................................................. 34

Figure 10: (Right and left plots) \( \text{O}_2 \) slant path density profiles for IPF/GOPR (blue) and ALGOM (red) for all occultations of faint star S0100. We see that GOPR/IPF provides unrealistic median (yellow...
Figure 11: \( \text{O}_2 \) slant path density profiles divided by air ecwmf for IPF/GOPR (right plot) and ALGOM (left plot) for all occultations of faint star S0002. We see that the dispersion is smaller for ALGOM than for GOPR/IPF by about 25-30%. The strong bias at low altitudes has disappeared. 35

Figure 12: \( \text{H}_2\text{O} \) slant path density profile for all occultations of star S0002 for ALGOM (red) and GOPR/IPF (blue). We can see on the zoomed box that the dispersion is smaller in ALGOM than in GOPR/IPF. 36

Figure 13: Upper plot: ALGOM retrieved PSF width as a function of orbit number for all occultations of S0004. We clearly see periodical fluctuations due to the fact that S0004 is a double star, double image which total PSF changes with the variable inclination of the line joining the two stars with the horizontal (spectral direction). These fluctuations are not taken into account in GOPR/IPF. Lower plot: \( \text{O}_2 \) slant path density profiles divided by air ecwmf for IPF/GOPR (right plot) and ALGOM (left plot) for all occultations of faint star S0004. We see that the dispersion is smaller for ALGOM than for GOPR/IPF by about 40%. 36

Figure 14: Same as Figure 51, see this figure for a description. We can see the improvement of using model transmission LUTs computed from GOMOS external model instead of climatological model. GOPR/IPF (not shown here) uses the LUTs computed from climatological model. 37

Figure 15: ALGOM retrieved slant path density \( \text{O}_2 \) profile whit the error bar correction factor on the measured transmissions (left) and without the error bar correction factor (right). The error bar correction factor allows retrieving more realistic error bars at low altitude. 37

Figure 16: \( \text{O}_2 \) slant path density profiles divided by air ecwmf for IPF/GOPR (right plot) and ALGOM (left plot) for all occultations of faint star S0001. Green circled: We can see that above 35 km, the median is closer to the expected value of 0.209 for GOPR/IPF than for ALGOM. 38

Figure 17: Same as Figure 16 for star S0002. As for S0001, we can see that above 35 km, the median is closer to the expected value of 0.209 for GOPR/IPF than for ALGOM. 39

Figure 18: Same as Figure 16 for star S0016. As for S0001, we can see that above 35 km, the median is closer to the expected value of 0.209 for GOPR/IPF than for ALGOM. 39

Figure 19: Same as Figure 16 for star S0063. As for S0001, we can see that above 35 km, the median is closer to the expected value of 0.209 for GOPR/IPF than for ALGOM. 39

Figure 20: Same as Figure 16 for star S0063. In this case IPF/GOPR shows between 35 and 65 km a positive bias compared to the expected value while ALGOM shows a negative bias at altitudes > 35 km. 40

Figure 21: Same as Figure 16 for star S0013. Contrary to all other star tested, S0013 shows a very good agreement between \( \text{O}_2 \) retrieved profile and air ecwmf because we see that the ratio is close to the expected value of 0.209. GOPR/IPF shows a positive bias at altitudes > 35 km. Note that below 15 km the deviation in ALGOM of the median and percentile toward small values is probably due to the occultation in PCD_IILLUM=4 in which the star is lost at low altitude. Star 13 probes high north and south latitudes, while the other do not or less. 40

Figure 22: Upper plot: black: measured transmission spectrum at 81.27 km in SPB1 extracted from the GOM_EXT product. Red: fitted model transmission spectrum by ALGOM processing. Blue: fitted model transmission spectrum by IPF/GOPR processing (extracted from the GOM_EXT product). Lower plot: red and blue: residuals for respectively ALGOM and IPF/GOPR processing. We see that at 81 km the measured spectrum is dominated by the noise, even for a bright star like Sirius. The gap in the residual spectrum between 765 nm and 766 nm corresponds to pixels which have been flagged. These pixels are not used by the Levenberg Marquardt. 44
Figure 23: Same as Figure 22 for a measurement at 25.25 km of tangent altitude. At this altitude, the O$_2$ spectral lines are very clear. Note that for the pixels with O$_2$ lines, the residual has higher amplitude for IPF/GOPR than for ALGOM. This is a consequence of the overestimation of the static spectral PSF in IPF/GOPR (see Figure 24).

Figure 24: Same as Figure 23 but zoomed in the spectral dimension between 762 nm and 765 nm. We see that the IPF/GOPR modelled transmission lines are smoother too much. Indeed, the PSF for IPF/GOPR is imposed and is over-estimated as explained in section 2.2.2 whereas ALGOM PSF width is a free parameter of the Levenberg-Marquardt. The chi2 is considerably reduced with ALGOM from 6.01 to 2.31.

Figure 25: Same as Figure 22 for a measurement at 6.76 km of tangent altitude. We note that at this very low altitude: 1. many lines between 760 and 765 nm are saturated. 2. The S/N is lower than at 25 km. 3. The IPF/GOPR fit failed. It is probably due to a problem in the IPF/GOPR (verified for both IPF and GOPR) for which the fit fails when transmission spectrum has at least one negative value (because it involves computing the log of a negative value which have probably not been taken into account).

Figure 26: Upper plots: log of O$_2$ slant path density profiles for S001 retrieved by ALGOM (red) and by GOPR (blue). The left and right plots are the same but on the left plot IPF/GOPR points are plotted over ALGOM points while on the right plot ALGOM points are plotted over IPF/GOPR points. Note also that the left plot show the medians while the right plot show the percentiles (16%, 84%). Lower plot: relative difference of the O$_2$ slant path density profiles between ALGOM and IPF/GOPR. The relative difference profile shows a negative bias up to -25% at 40 km. An attempt to explain this bias is provided in section 8.1.9. We see that the GOPR density profiles systematically fail below about 13 km (vertical stripes) due to a problem in the processor (computation of the logarithm of negative transmission values). Those points explain the behaviour of the relative difference plot below 13 km (huge positive bias). Those points are flagged by IPF/GOPR.

Figure 27: Error bars (absolute) on the logarithm of the slant path density profiles. There is a bug in the version 6 of the IPF level2 products: the error bars on slant path densities have been wrongly coded in the product and it is impossible to recompute them for SPB. As a consequence, this plot has been done by using products generated with the version 5 of IPF instead of version 6. We see that between 40 and 70 km, the error bars are smaller by about 20% for ALGOM.

Figure 28: The slant path density profiles of ALGOM (left) and IPF/GOPR (right) are compared to the ECMWF air density profiles. The expected value is 0.20948, the mixing ratio of O$_2$. It is not clear if ALGOM is closer to this value than IPF/GOPR. Indeed, above 40 km, it seems that there is a negative bias for ALGOM. We note also that the dispersion above 40 km seems lower for ALGOM than for IPF/GOPR. The divergence for IPF/GOPR below 13 km is due to the fit failure (see Figure 25).

Figure 29: Chi2 profile for ALGOM (red) and IPF/GOPR (blue). The chi2 is always smaller for ALGOM and stays >1. This is mostly due to a better estimate of the PSF by ALGOM, leading to smaller residuals as shown in Figure 23 and Figure 24. Another contribution to this lower chi2 can be the better estimate of the wavelength assignment (see Figure 31). Below 13 km, IPF/GOPR fit fails (see Figure 25).

Figure 30: Lower plot. Left: profiles of the retrieved global spectral PSF sigma (median in black). Right: associated error bars. Upper plot: width of the global PSF at 24 km altitude as a function of the orbit number (2002 -> 2012). We see on the upper plot that the retrieved PSF sigma at 24 km is about 0.024 nm at the beginning of the mission and slightly decreases to about 0.020 nm at the end of the mission. There is no seasonal variation of the PSF sigma. We see on the left lower plot that the PSF width starts to increase below 25 km due to increased scintillations. However below about 10 km, the quality of the sigma PSF retrieval decreases probably because of lower SNR and also because...
most lines become saturated (see Figure 25). On the same plot we can notice that the PSF width is very slightly lower at 25 km than above 40 km.

Figure 31: Same as Figure 30 but for the wavelength assignment shift. On the upper plot, the nominal wavelength assignment (extracted from the GOM_EXT product) has been plotted in black. This nominal wavelength assignment is used by IPF/GOPR. ALGOM improves this wavelength assignment by adding a retrieved shift to this nominal value. We see on the upper plot that the retrieved ALGOM wavelength assignment has seasonal variations (kind of ‘V’ shape). The strong dispersion of the points on the wl shift profile plot (lower left) is due to these seasonal variations. If we look carefully at all individual occultations, it is possible to see that the wavelength shift slightly increases below 15 km.

Figure 32: Upper left: profile of the retrieved height baseline coefficient. Upper right: associated error bars. Lower left: profile of the retrieved slope baseline coefficient. Upper right: associated error bars.

Figure 33: Same as Figure 26 but for a dataset made of star S0063. We find the same negative bias between ALGOM and GOPR/IPF than for S0001 dataset. Below 10-15km, ALGOM shows some outliers which are due to the low SNR at low altitude. Note the very small spread of values in the range 20-30 km of altitude, compared to star S0001 (figure 26). This is because star S0063 probes only intertropical latitudes while S0001 probes most latitudes, with a much larger natural latitude variations.

Figure 34: Same as Figure 26 but for a dataset made of star S0041. Note the very small spread of values in the range 20-30 km of altitude, compared to star S0001 (figure 26). Same comment as in Figure 33.

Figure 35: Same as Figure 26 but for a dataset made of star S0029. GOPR/IPF profiles fail systematically below about 23 km.

Figure 36: Same as Figure 26 but for a dataset made of star S0157. IPF/GOPR profiles are most erratic at all altitudes.

Figure 37: Retrieved SPB1 PSFs sigma profiles for several stars (coloured lines) and mean of all stars (black line) for early mission (left) and late mission (right). The mean curves are stored and can be used when the ALGOM configuration is set on ‘FIXED’ PSF width in the Levenberg Marquardt. Interpolation in time and in altitude in these LUTs will be used.

Figure 38: Retrieved slant path density $O_2$ profiles divided by air ecmwf with PSF fixed (red) or PSF free (blue) in the Levenberg Marquardt for all occultations of star S0001 (upper), S0063 (lower). Left is red plotted over blue, right is blue plotted over left. Median is visible on left plot (thick line) and percentiles on right plot (dotted line), Colors for median and percentile is yellow for PSF fixed and black for PSF free. For the two stars, the difference between PSF free and PSF fixed is negligible except a small reduction of the dispersion for PSF fixed above 60 km.

Figure 39: Retrieved SPB1 wavelength shift profiles for several stars (colored lines) and mean of all stars (black line). The mean curve is stored and can be used when the ALGOM configuration is set on ‘FIXED’ WL in the Levenberg Marquardt. In this case, a retrieval of wavelength shift at 25 km with ‘FREE’ WL is performed, then the LUT is applied in order to find the wl shift at other altitudes.

Figure 40: Retrieved slant path density $O_2$ profiles divided by air ecmwf with 25km+LUT retrieved Wavelength shift (red) or free wavelength shift (blue) in the Levenberg Marquardt for all occultations of star S0001 (upper), S0002 (middle) and S0029 (lower). Median is visible on left plot (thick line) and percentiles on right plot (dotted line), Colors for median and percentile is yellow for wl 25km+LUT and black for wl free. Left is red plotted over blue, right is blue plotted over left. The difference between blue and red points is negligible for the bright stars S0001 and S0002. There is a small
difference for S0029 above 50 km but it is not obvious that the profiles are improved in one case compared to the other................................................................. 62
Figure 41: O₂ model transmission LUT at tangent height altitude of 20 km for different versions (2004, 2008 and 2012) of HITRAN smoothed to a resolving power of 35000@764nm, for HITRAN 2012 infinite resolution and for GOPR LUT (black). We can see that the sampling of the GOPR LUT is too low to preserve the shape of the lines (lines become jaggy). We see also that the wavelength of the lines is not the same between GOPR LUT and HITRAN LUTs. GOPR LUT lines wavelength is generally smaller than HITRAN by about 0.01nm but this is not always the case like for example the first line of this plot, near 764.03nm, which is located at the same wavelength for HITRAN and GOPR LUTs................................................................. 64
Figure 42: Comparison of O₂ slant path profiles from S0001 dataset between using the GOPR transmission model LUTs vs. the HITRAN 2004 recomputed transmission model LUTs. There is a negative bias above 30 km reaching about -5% at 40 km................................................................. 65
Figure 43: Comparison of Chi² profiles for O₂ retrieval of Figure 42.................................................. 66
Figure 44: Same as Figure 42 but for HITRAN 2004 recomputed transmission LUTs vs. HITRAN 2008 recomputed transmission LUTs. The differences are small (<1%)................................................................. 66
Figure 45: Same as Figure 42 but for HITRAN 2004 recomputed transmission LUTs vs. HITRAN 2012 recomputed transmission LUTs. The differences are small (<1%)................................................................. 67
Figure 46: Comparison of two runs of O₂ retrieval on S0001 dataset: run 1 = without taking into account the CIA (blue in upper plots), run 2 = taking into account the CIA (red in upper plots). The lower plot shows the relative difference. As expected, the impact is only below 20 km where the ‘with CIA’ run retrieves less O₂ than the ‘without CIA’ run.................................................. 68
Figure 47: Chi² profile and chi² relative difference of the runs presented in Figure 46. When taking into account the CIA, the chi² is slightly reduced below 20 km (up to -5% near 10-15km). .............. 69
Figure 48: Comparison of the PSF width (sigma) between the two runs presented in Figure 46. The lower right plot shows that when taking into account the CIA, the PSF width is slightly decreased (-2%) at altitudes (> 20 km) where we would not expect any influence of the CIA. The impact below 20 km is bigger: there is a negative bias up to -20% at 5 km................................................................. 69
Figure 49: Left: O₂ vs. ecmwf profile for a run without taking into account CIA. Right: same as left plot but for a run taking into account the CIA. We see that the break of the curve around 10-15km on the left plot has disappeared on the right plot. This result tends to prove that the CIA correction helps removing a bias in the O₂ profiles at low altitude.................................................. 70
Figure 50: Upper plots: O₂ slant path density profiles obtain either by taking the good (lat/season dependant) climatological model LUT (red) or by forcing the transmission model to US STANDARD (blue). Lower plot: relative difference profile. The colors correspond to the climatological conditions (see the legend in the upper right corner of the plot). We see that the retrieved O₂ profile is very sensitive to the choice of the climatological transmission LUT as different climatological profile can lead to O₂ slant path density differences of up to +50% around 40 km.................................................. 71
Figure 51: O₂ slant path density profiles divided by Gomos external model air density for a dataset made of 191 occultations of star S0001. Note that this star covers a wide range of latitude which is good for the current test. Blue: the model transmission LUTs comes from climatological model. Red: the model transmission LUTs come from GOMOS External atmosphere model (ecmwf + msis). Median is visible on left plot (thick line) and percentiles on right plot (dotted line), Colors for median and percentile is yellow for GOMOS EXTERNAL and black for CLIMATOLOGY. Using the model transmission LUTs from GOMOS external model seems to improve the results because: - the dispersion is reduced (check percentile on right plot) between 40 and 70 km. – the negative bias in the median between 40 and 70 km is reduced.................................................. 72
Figure 52: Relative difference of O2 retrieved profiles between ALGOM and IPF/GOPR, for S0001 dataset, for different values of the PSF sigma fixed in the Levenberg Marquardt (only above 20 km, below 20 km the PSF sigma is let free, which explains the break at 20 km). We see that the negative bias present for small sigma PSF values tends to disappear when the sigma PSF increases.

Figure 53: Same as Figure 52 but for the Chi2 profiles. Blue is IPF/GOPR, red is ALGOM. Above 20 km, where the PSF has been fixed, the chi2 is almost equivalent between IPF/GOPR and ALGOM for PSF sigma=0.045 nm which signify that the global PSF sigma used by IPF/GOPR is probably close to 0.045 nm.

Figure 54: Transmission spectrum at 37 km for a Sirius occultation. The correction factor is computed as the ratio between the standard deviation of the pixels of the continuum (red circle) by the median of the error bars of the same pixels.

Figure 55: Upper plots: retrieved O2 slant path density profiles with associated error bars in green for the case without applying the correction factor (left) and the case applying the correction factor (right). Lower plots are a zoom for altitude < 20 km. We see that with the correction factor applied, the error bars at low altitudes are increased and more realistic than without the correction factor.

Figure 56: Upper left: chi2 profiles for the case applying the correction factor (blue) and the case not applying the correction factor (red). Upper right: same as upper left with blue overplotted on red. Lower left: retrieved O2 density error bars. Lower right: same as lower left with blue overplotted on red. We can see that the chi2 is reduced and closer to 1 when the correction factor is applied. The error bars below 20 km are increased when correction factor is applied.

Figure 57: Sfa azimuth profile for R32522 S0016 in pcd_iilum=4 illumination conditions. We see that at about 10 km, there is a break of the curve which is relevant of a false star tracking: the star tracker starts to follow something else than the star. A method is proposed in [RD-5] for automatically finding the presence of a break in the sfa elevation angle curve.

Figure 58: Left: O2 slant path density profile for the case of Figure 57 (R32522 S0016) for ALGOM (red) and GOPR/IPF (blue). We see on the ALGOM profile that the retrieval becomes very bad below about 10 km which is due to the fact that the star is not tracked anymore. This is not reflected in the error bars (in black) which are still very small. GOPR/IPF profile fails below about 17 km (known problem already mentioned in this TN, not related with star tracking). Right: error bars of retrieved ALGOM O2 density profiles.

Figure 59: ALGOM retrieved O2 slant path profiles for star S013 (left) and star S016 (right). Colors refer to the illumination condition of the occultation. We can see in both plots that many profiles in pcd_iilum=4 (red) are bad (too small retrieved values) at altitudes < 10 km. It seems that for these profiles, the star is lost (the SATU starts tracking something else than the star) at the altitude where the density value starts to be biased.

Figure 60: Upper left: ALGOM O2 chi2 profile for dataset made of 579 products of star S0063 close in time (between R42000 and R44000). We see that there are two groups of points: first group (colored in red) with small chi2 value, second group (colored in blue) with much higher chi2 below 30 km. For the red points the DSA chosen for DC correction in the Level-1B processing is very close to the processed occultation (typically less than 5 minutes) while for the blue points it is much further (typically about 1h30, one orbit away). Upper right: slant path O2 density profiles at altitudes < 13 km for the same dataset. We see that the blue points (high chi2) are more dispersed than the red points. Lower right: errors bars of the slant path O2 density profiles at altitudes < 13 km. We see that the higher dispersion of the blue points in the O2 profile is not reflected in the error bars which are roughly the same for blue and red points. Exemps of transmission spectrum at low altitude for the two groups of points are provided in Figure 61.
Figure 61: Exemple of two measured transmission spectrum at about 8 km for a case when the dark charge correction is good (upper) and a case when the dark charge correction is bad (lower). In the upper plot, the DSA used for DC correction is very close in time from the occultation. In the lower plot the DSA used for the DC correction is far from the occultation. The upper plot corresponds to an occultation in the red group of Figure 60 while the lower plot corresponds to an occultation in the blue group. ................................................................................................................. 80

Figure 62: Upper left: chi2 profile for ALGOM O2 spectral inversion for a dataset of 579 occultations of S063 between R42000 and R44000. The Level-1B processing comes from IPF-6 for the blue points and GOPR80da for the red points. Upper right: same as upper left with blue points plotted over red points. Lower: relative difference between blue and red points (reference is blue=IPF6). In the lower plot, below 30 km we see 3 groups of points: 1rst group is close to 0% which is relevant of points for which the same DSA was used in IPF6 and in GOPR80da. Therefore both IPF6 and GOPR80da could choose the closest available DSA. 2nd group is the one that goes around -40% below 20 km (with a high dispersion). For this group, the DSA used by GOPR80da is closer to the occultation than the DSA used by IPF6, however this closest available DSA is still at about 15-20 min from the occultation. The 3rd group is the one that goes at about -75% below 20km. For this group of point we go from a very far DSA (about 1h30) used by IPF6 to a very close DSA (<5min) used by GOPR80da. The slant path densities profiles are presented in Figure 63................................................................................................................. 81

Figure 63: O2 slant path density profiles at altitude < 17 km for the dataset of 579 occultations of S063 between R42000 and R44000. The Level-1B processing comes from IPF-6 for the blue points and GOPR80da for the red points. Upper right: same as upper left with blue points plotted over red points, 5% 95% percentiles are in thick dotted line (black for IPF6, Yellow for GOPR80da). Median is visible on left plot (thick line) and percentiles on right plot (dotted line), Colors for median and percentile is yellow for 80da and black for IPF6. We see a reduction of the dispersion especially below 10 km for GOPR 80da compared to IPF6. This is due to the fact that for some occultations GOPR80da could find a closest DSA than IPF6. ................................................................................................................. 82

Figure 64: Upper plot: black: measured transmission spectrum at 60.35 km in the last pixels (high wavelengths) of SPA2 extracted from the GOM_EXT product. Red: fitted model transmission spectrum by ALGOM processing. Blue: fitted SPA model transmission spectrum by IPF/GOPR processing (extracted from the GOM_EXT product). O2 is not taken into account in this fit. Lower plot: red and blue: residuals for respectively ALGOM and IPF/GOPR processing. The O2 B absorption line can be seen at 687-688 nm but at this altitude the SNR is still very low, even for a bright star like Sirius. ................................................................................................................. 83

Figure 65: Same as Figure 64 at 22.15 km. The O2B absorption line is very clear seen between 686 and 690 nm............................................................................................................. 84

Figure 66: Retrieved SPA2 PSF sigma for all occultation of Sirius from the inversion using the O2B absorption band. We build a look up table (thick orange line) in order to be able fix the PSF width in LM............................................................................................................. 84

Figure 67: Upper: retrieved O2 slant path profile for all occultations of S0002 using SPB1/O2A absorption band only (red) and SPA2/O2B absorption band only (blue). Lower: retrieved O2 slant path profile divided by air ecmwf. Median is visible on left plot (thick line) and percentiles on right plot (dotted line), Colors for median and percentile is yellow for SPB1/O2A and black for SPA2/O2B........... 85

Figure 68: Same as Figure 67 for S0029. We see that the median of the ratio O2/air starts to diverge above 35 km............................................................................................................. 85

Figure 69: Same as Figure 64 for combined O2A + O2B retrieval at 25.35 km. .............................................. 86

Figure 70: Upper plots: O2 slant path density profiles obtain with O2A (SPB1) absorption band only (red) and with O2A (SPB1) + O2B (SPA2) absorption bands (blue). Lower plot: relative difference
profile. Dataset is about 200 occultations of star S0002. We do not see obvious differences between the two kinds of retrieval. The comparison with ecmwf is presented in Figure 71. Figure 71: retrieved O$_2$ slant path density profile divided by air ecmwf obtained with O$_2$A (SPB1) absorption band only (red) and with O$_2$A (SPB1) + O$_2$B (SPA2) absorption bands (blue). Median is visible on left plot (thick line) and percentiles on right plot (dotted line). Colors for median and percentile is yellow for SPB1/O$_2$A and black for O$_2$A (SPB1) + O$_2$B (SPA2). Dataset is about 200 occultations of star S0002. We do not see obvious differences between the two kinds of retrieval. Figure 72: Error bars for the O$_2$ profiles presented in Figure 70. We notice a small reduction of the error bars above 40 km when using both O$_2$A and O$_2$B absorption bands. Figure 73: Same as Figure 70 for star S0029. We do not see obvious differences between the two kinds of retrieval. Figure 74: Same as Figure 71 for star S0029. We do not see obvious differences between the two kinds of retrieval. Figure 75: Upper plot: black: measured transmission spectrum at tangent altitude = 49.12 km in SPB2 extracted from the GOM_EXT product. Red: fitted model transmission spectrum by ALGOM processing. Blue: fitted model transmission spectrum by IPF/GOPR processing (extracted from the GOM_EXT product). Lower plot: red and blue: residuals for respectively ALGOM and IPF/GOPR processing. We see that at 49 km the measured spectrum is dominated by the noise, even for a bright star like Sirius. Figure 76: Same as Figure 75 for a measurement at 24.16 km of tangent altitude. At this altitude, the measured transmission spectrum is severely contaminated by residual iPRNU structures with amplitude higher than the H$_2$O lines. This shows that the iPRNU correction is insufficient at this altitude even though it still performs better than the nominal PRNU correction (used in IPF v5). Figure 77: Same as Figure 75 for a measurement at 7.34 km of tangent altitude. At this altitude, the measured transmission spectrum shows some strong H$_2$O absorption lines. The residual is very close between ALGOM and GOPR/IPF. Figure 78: Upper plots: log H$_2$O slant path density profiles retrieved by ALGOM (red) and by GOPR (blue). The left and right plots are the same except on the left plot IPF/GOPR points are plotted over ALGOM points while on the right plot ALGOM points are plotted over IPF/GOPR points. Note also that the left plot shows the medians while the right plot shows the percentiles (16%, 84%). Lower plot: relative difference of the H$_2$O slant path density profiles between ALGOM and IPF/GOPR. The relative difference profile shows a negative bias up of about -15% above 15 km. We see that some GOPR density profiles fail below about 10 km (blue vertical stripes) due to a problem in the processor. This problem exists also for O$_2$ retrieval (see Figure 26). It seems that there are more outliers in ALGOM above 15 km (and especially above 40 km) than for IPF/GOPR. This is due to the fact that the PSF width is kept free for ALGOM (for this test) while it is fixed (to a rather correct estimation) for IPF/GOPR. Figure 97 presents the same plots for the ALGOM-delivered H$_2$O dataset adding a median plotted on the lower plot. Figure 79: Error bars (absolute) on the logarithm of the slant path density profiles. There is a bug in the version 6 of the IPF level2 products: the error bars on slant path densities have been wrongly stored and it is impossible to recompute them for SPB. This is why only ALGOM error bars (in red) appear in this plot. Figure 80: Chi2 profile for ALGOM (red) and IPF/GOPR (blue). Above 10 km, ALGOM and IPF/GOPR show very close chi2. Below 10 km, some IPF/GOPR fit fails as explain above in this document.
one for SPB1 (see O₂ retrieval, Figure 30). The decrease of the PSF width when the altitude increases is probably a bias introduced by the values for which the retrieved PSF width = 0.001 nm (minimum allowed PSF value by the LM). The higher is the altitude the more of these values there is because absorption features become smaller when altitude increases. These points should be eliminated because they probably correspond to unsuccessful fits.

Figure 82: Same as Figure 81 but for the wavelength assignment shift. On the upper plot, the nominal wavelength assignment (extracted from the GOM_EXT product) has been plotted in black. This nominal wavelength assignment is used by IPF/GOPR. ALGOM improves the wavelength assignment by adding a possible shift to this nominal value. We see on the upper plot that the wavelength assignment has seasonal variations (same kind of 'V' shape than for SPB1 (see Figure 31)). The strong dispersion of the points on the wl shift profile plot (lower left) is due to these seasonal variations.

Figure 83: Upper left: profile of the retrieved height baseline coefficient. Upper right: associated error bars. Lower left: profile of the retrieved slope baseline coefficient. Upper right: associated error bars.

Figure 84: Same as Figure 37 for SPB2/H₂O. Above 35 km the SNR of the spectrum is two small giving over-estimated retrieved PSF sigma. Consequently we take a constant value above 35 km for our LUTs.

Figure 85: Retrieved slant path density H₂O profiles with fixed PSF width (red) or free PSF width (blue) in the Levenberg Marquardt for all occultations of star S0001 (upper), S0014 (middle) and S0063 (lower). Left is red plotted over blue, right is blue plotted over left. Letting the PSF free provides many wrong profiles above 15 km for the three stars. The cause is the iPRNU structures which are fitted by the PSF instead of the H₂O absorption lines. Fixing the PSF to the expected LUT value removes a lot of these bad profiles. We can see however that even when fixing the PSF there are still a few profiles that are biased like for example S0014 above 35 km (see Figure 96 for a dedicated figure on this feature). It is important to note that these results are not an improvement compared to GOPR/IPF which uses also a fixed PSF width.

Figure 86: Same as Figure 39 but for the wavelength assignment shift. On the upper plot, the nominal wavelength assignment (extracted from the GOM_EXT product) has been plotted in black. This nominal wavelength assignment is used by IPF/GOPR. ALGOM improves the wavelength assignment by adding a possible shift to this nominal value. We see on the upper plot that the wavelength assignment has seasonal variations (same kind of 'V' shape than for SPB1 (see Figure 31)). The strong dispersion of the points on the wl shift profile plot (lower left) is due to these seasonal variations.

Figure 87: Upper: retrieved H₂O slant path density profiles with 25km retrieved + LUT wavelength shift (red) or free wavelength shift (blue) for all occultations of star S0063. Left is red plotted over blue, right is blue plotted over red. Lower: relative difference profile. Above 40 km the 25km+LUT configuration seems to deliver less H₂O and the median (upper left plot) seems in better continuity with lower altitudes. Below 40 km the relative difference seems also slightly smaller (between -0% and -10%) for the 25km+LUT configuration.

Figure 88: Upper left: ALGOM retrieved wavelength assignment as a function of orbit number at 23.5 km for pixel 200 of SPB1 (blue) and pixel 200 of SPB2 (red) obtained respectively from ALGOM O₂ inversion and ALGOM H₂O inversion. The dataset used is all occultations of star S0014. Lower left: difference between SPB2 and SPB1 retrieved wavelength assignment. Right: scatter plot of SPB2 vs SPB1 wavelength assignment. We see a trend that starts at about orbit 13000 + some seasonal variations. The color refers to the orbit number (same color code as in lower left plot). We see that the scatter plot follows a linear model but that the intercept of this linear model varies with time. This variation of the intercept is the trend that we see in the lower left plot.

Figure 89: Same as Figure 88 for star S0001: results are very similar to S0014.

Figure 90: Same as Figure 88 with the model $wl_{SPB2} = f(wl_{SPB2})$ overplotted on the right plot. The model $f()$ is linear with the intercept varying linearly with time for orbit $>13000$. The variation of the intercept is represented in the lower left plot by a black curve.

Figure 91: Upper plots: H₂O slant path density profiles for all occultations of star S0063. Blue: wavelength assignment was fixed to a value computed using the SPB1 -> SPB2 wl model, red:
wavelength assignment is directly retrieved from the SPB2 H₂O inversion. We see that there is almost no difference between the two kinds of retrieval. Lower plot: relative difference profile between blue and red points. ............................................................................................................................. 106

Figure 92: Upper plots: H₂O slant path density profiles obtained with 2 different model transmission LUTs: GOPR LUT in blue, HITRAN 2004 in red. Lower plot: relative difference profile. Contrary to O₂ results, there is no negative bias. However the dispersion is still large. ........................................... 107

Figure 93: Same as Figure 92 but comparing HITRAN 2004 and HITRAN 2008 model transmission LUTs. There is not a big impact. ............................................................................................................................. 108

Figure 94: Same as Figure 92 but comparing HITRAN 2004 and HITRAN 2012 model transmission LUTs. The use of HITRAN 2012 leads to a bias of about (-2, -3 %) which is probably the result of some changes of H₂O spectral structures (or absorption lines temperature coefficients) in the latest version of HITRAN. ............................................................................................................................. 109

Figure 95: Upper plots: H₂O slant path density profiles obtained either by taking the good (lat/season) climatological model LUT (red) or by forcing the transmission model to US STANDARD (blue). Lower plot: relative difference profile. The colors correspond to the Climatological conditions (see the legend in the upper right corner of the plot). As for O₂ retrieval, H₂O retrieval is also very sensitive to the choice of the climatological model. Here there is a bias of up to -20% at low altitude. ............................................................................................................................. 110

Figure 96: Left: slant path H₂O profiles for all S0014 occultations. Right: associated error bar profile. The color code is related to the orbit number (black for early mission to red for late mission). We see that the profiles in the green circle have a positive bias between 30 and 50 km. Their associated error bars are very small compared to the error bars of the non biased profiles. The biased profiles come probably from a bad iPRNU correction. ............................................................................................................................. 111

Figure 97: Upper plots: log H₂O slant path density profiles retrieved by ALGOM (red) and by GOPR (blue). The dataset used is the ALGOM-delivered H₂O dataset, composed of the 8 brightest NIR stars (S004 excluded). ALGOM processor has been configured with the parameters of Table 5. The left and right plots are the same except on the left plot IPF/GOPR points are plotted over ALGOM points while on the right plot ALGOM points are plotted over IPF/GOPR points. Note also that the left plot shows the medians while the right plot shows the percentiles (16%, 84%). Lower plot: relative difference of the H₂O slant path density profiles between ALGOM and IPF/GOPR. The relative difference profile shows a negative bias with median of -10% above 15 km................................................................. 112

List Of Tables

Table 1: Applicable Document ........................................................................................................... 15
Table 2 : Reference Document ........................................................................................................ 16
Table 3: main characteristics of GOMOS SPB1 and SPB2 spectrometers. .................................. 18
Table 4: default configuration for the parameters of the levenberg-marquardt. The Allowed range is the range of tested values constrained by the Levenberg Marquardt. It never uses values which are outside of this range................................................................. 26
Table 5: Configuration parameters delivering the best profiles. These parameters have been used for generating the ALGOM sample datasets delivered to ESA................................................................. 32
1 Introduction

1.1 Scope of the document

This document is a technical note presenting the results of the activities performed in the frame of WP2 (development of an improved H₂O algorithm) and WP3 (development of a new O₂ algorithm) of the ALGOM project.

1.2 Structure of the document

This document is split into several chapters:

- This chapter introduces the document;
- Chapter 2 - (common for O₂ and H₂O) - Gives a description of the current algorithm;
- Chapter 3 - (common for O₂ and H₂O) - Presents the new ‘ALGOM’ algorithm;
- Chapter 4 - (common for O₂ and H₂O) - Presents the generation of the ALGOM datasets;
- Chapter 5 - (common for O₂ and H₂O) - Summarizes the improvements brought by ALGOM for O₂ and H₂O retrieval;
- Chapter 6 - (common for O₂ and H₂O) - Presents possible further improvements (not covered by this project);
- Chapter 7 - (common for O₂ and H₂O) - Draws the general conclusions of the ALGOM O₂/H₂O project;
- Chapter 8 - (separated in an O₂ part and a H₂O part) - Presents in detail the results of all the runs performed with the new algorithms on various datasets and with various configurations of the ALGOM processor.

1.3 Applicable and Reference Documents

1.3.1 Applicable Document

The following table presents the Applicable Document (AD).

Table 1: Applicable Document

<table>
<thead>
<tr>
<th>ID</th>
<th>Title</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>[AD-1]</td>
<td>ALGOM: Improvements for GOMOS O2 and H2O profiles retrieval – Contract</td>
<td>Contract n° 0166L68540</td>
</tr>
</tbody>
</table>

1.3.2 Reference Document

The following table presents the Reference Document (RD).
Table 2: Reference Document

<table>
<thead>
<tr>
<th>ID</th>
<th>Title</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>RD-1</td>
<td>GOMOS Level 2 algorithm evolution studies ALGOM</td>
<td>FMI-2014-ALGOM-1</td>
</tr>
<tr>
<td>RD-2</td>
<td>Rétro ingénierie du traitement des données du spectromètre B1 (O2)</td>
<td>Author: Jean-Luc Vergely</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ref: ACRI/Service d’Aeronomie internal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Technical Note</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Date: 2003/11/28</td>
</tr>
<tr>
<td>RD-3</td>
<td>ALGOM – Calcul des transmissions O2 et CIA pour le spectromètre B1 de GOMOS</td>
<td>Author: Stéphane Ferron</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ref: ACRI/LATMOS internal Technical Note</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Date: 2014/10/16</td>
</tr>
<tr>
<td>RD-4</td>
<td>False Star Tracking</td>
<td>Author: Gilbert Barrot</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ref: GOMOS QWG 25 presentation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(qwg_25_gb_check_star_tracking.ppt)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Date: 2011/09/26</td>
</tr>
<tr>
<td>RD-5</td>
<td>Ozone Bias and dark charge correction</td>
<td>Author: Laurent Blanot</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ref: Technical note in the frame of ALGOM CCN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Date: JAN 2016.</td>
</tr>
<tr>
<td>RD-6</td>
<td>iPRNU correction quality investigation</td>
<td>Author: Laurent Blanot</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ref: GOMOS QWG N2 presentation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(LBL_QWG_N2_iPRNU_correction_quality.ppt)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Date: 26 Nov 2014</td>
</tr>
</tbody>
</table>

1.4 Acronyms

The definition of the acronyms used in this document is provided below:

DC: Dark Charge
DSA: Dark Sky Area
GOMOS: Global Ozone Monitoring by Occultation of Stars
iPRNU: intra-Pixel Response Non Uniformity
SAA: South Atlantic Anomaly
SFA: Steering Front Assembly
SPB: Spectrometer B
SPA: Spectrometer A

© 2016 ACRI-ST
2 Current Algorithm

2.1 Description

The current algorithm used in IPF 6 (and also GOPR 7 which is equivalent) for O₂ (Spectrometer B1) and H₂O (Spectrometer B2) spectral inversion is not the same as the algorithm used for spectrometer A retrieval (O₃, NO₂, NO₃, aerosols). The spectrometer B algorithm uses a method developed at Service d’Aéronomie (now LATMOS) by Jean-Loup Bertaux. This algorithm is equivalent for O₂ and H₂O. It relies on a comparison between the measured transmission spectrum (Tmes) and model transmission spectra LUTs (Tmod). An example of Tmes and Tmod is shown in Figure 1. For each measurement, a scatter plot of log(Tmes) vs. log(Tmod) is performed for various model altitudes. When the slope of the linear fit of the scatter plot is 1, then the O₂ (or H₂O) line density of the model at this altitude is declared to be the good one measured. Such a scatter plot is represented in Figure 2. Once the line density profile has been retrieved, the same vertical inversion as for Spectrometer A is performed in order to retrieve local densities.

Figure 1: O₂ measured (red) and model (black) transmission spectrum in SPB1. The tangent point altitude of the measured transmission is 14.5 km. The baseline of the measured transmission is lower than 1 because of aerosol and Rayleigh scattering effects. Those effects are not considered in the represented model transmission spectrum.
The current Spectrometer B retrieval suffers from several limitations.
2.2.1 Wavelength assignment

The variation with time (or orbit number) of the wavelength assignment of the spectrometer pixels (spectels) is imposed. It uses a model which has been computed by fitting a 2nd order polynomial of the wavelength shift vs. orbit number from star S0002 data only (see Figure 3). The wavelength shift of star S0002 shows only a global trend but no seasonal variation. This is not the case for many other stars which show strong seasonal variations as, for example, star S0001, represented in Figure 4. These are effects related to gradient of temperature of the whole GOMOS instrument, which changes the alignment between the star tracker unit (SATU) and the spectrometers entrance slit so it’s not caused by the star itself. But each star is observed in different seasons and the GOMOS temperature conditions are different (solar illumination, latitude of ENVISAT…). The fact that the seasonal variations are not taken into account in the current processor can be a limitation for the quality of the retrieval. Indeed, a sensitivity analysis (cf. GOMOS QWG AI 22.3) of the wavelength shift (forcing it to the three values -0.015, 0, +0.015nm)) has shown that O2 profiles may be impacted to up to 10% (see Figure 5).

![Figure 3: Wavelength assignment shift as a function of orbit number for spectrometer B1 computed from S0002 data by cross-correlation of a low altitude measured atmospheric O2 transmission spectrum vs. a reference spectrum. Measurement spectrum has been taken from the residual extinction product (GOM_EXT). In blue, the reference spectrum used is a measurement spectrum near orbit R03000. In green, the reference spectrum used is the model spectrum of the current RES_EXT product but associated with the wavelength assignment near orbit R03000. Both methods give nearly similar results. The blue data has been fitted by a 2nd order polynomial (red curve), which provides the ‘imposed’ wavelength assignment shift model used by IPF 6.01 (and GOPR 7) for the processing of all GOMOS occultations (applied to all stars and all seasons).](image-url)
Figure 4: Same as Figure 3 but for star S0001. We see strong seasonal variations of amplitude up to 0.05nm.

Figure 5: Relative difference of $O_2$ tangent line density profiles between: 1. wavelength assignment shift=0 nm (nominal value). 2. Wavelength assignment shift forced to +0.015 nm (= 1/3 pixel). Y-axis is altitude in km. The $O_2$ relative difference reaches about 10% between 40 and 60 km. This study confirms the importance of having a precise wavelength assignment determination for the retrieval of $O_2$ profiles. For this, the seasonal variation of the wavelength assignment shift must be taken into account.

2.2.2 Spectral Point Spread Function (PSF)

The spectral PSF can be decomposed in 2 parts:

- The Static PSF. This is the PSF for an ideal case where the star position on the CCD would be perfectly stable during a measurement;
The dynamic PSF. This is the contribution due to the displacement of the star on the CCD in the X (spectral) axis during one measurement. This can be computed from the Satu-X data. The displacement of the star is mainly due to scintillation effects in the atmosphere and happens only at low altitudes.

In the Current algorithm, the static PSF is read in an ADF (instrument physical configuration (IPC) product) and is thus the same for all occultations. The dynamic contribution of the PSF is added to the static one by using the Satu-X data, representative of the displacement of the star in the spectral direction (50 values per measurement). The resulting PSF (static + dynamic) is called the ‘global’ PSF.

A study led by J.L Vergely at Service d’Aéronomie (cf. [RD-2]) has shown that for SPB1, the width of the static PSF provided by Matra, which is stored in the IPC product is overestimated by about 30% compared to the PSF that he re-computed in the frame of his study (for altitudes above scintillations). The IPC PSF is a common PSF for SPB1 and SPB2. It has been decided (source: mails between G.B, R.F and J.L.B from 2003) that the IPC PSF would be calculated as a pure Gaussian LUT with a width equal to the average of the 6 SPB on-ground measured PSF width, plus an addition of 0.1 nm for taking into account possible micro-vibrations of the instrument. The SPB on-ground PSFs were measured by Matra in 1998 in thermal vacuum conditions (TVC) at 6 different wavelengths: 3 wavelengths in spectrometer B1 (756.0 nm, 766.5 nm and 769.9 nm) and 3 wavelengths in spectrometer B2 (930.0 nm, 942.0 nm, 769.9 nm). We can see on Figure 6 that the IPC PSF is larger than the measured PSF at 766.5nm and 769.9nm. It is however narrower than the measured PSF at 756 nm but the O$_2$ absorption lines occur only above 759 nm in SPB (as shown in Figure 1). Thus, we have here the confirmation that the SPB1 IPC PSF is overestimated, as shown in Jean-Luc Vergely study. In SPB2, the IPC PSF width is rather close to the 3 measured PSF and thus should not be considered as overestimated.
Figure 6: 3 coloured curves: spectral PSFs measured on ground in thermal vacuum conditions in 1998 for 3 different wavelengths covered by spectrometer B1. Black curve: The Spectral PSF stored in the IPC product (Gaussian). The measured PSFs at 766.5 nm and 769.9 nm have a very similar width whereas the PSF at 756 nm is larger. We note that the IPC PSF is larger than the measured PSFs at 765.5 nm and 769 nm but smaller than the 756.0 nm. There is no O₂ absorption lines in SPB1 for wavelengths below 759 nm (see Figure 1), thus a width of the PSF closer to the 766.5 and 769.9 nm values would have been more appropriate to set in the IPC product. This is what has been shown in [RD2].

Figure 7: Same as Figure 6 for spectrometer B2. The IPC PSF is the same as for SPB1. It is close to the measured PSF widths.

As for the wavelength assignment, a wrong estimation of the PSF width can be a problem for the retrieval of O₂ and H₂O profiles. Indeed, the absorption lines of these species are narrow
and thus the forward model is sensitive to the convolution of the transmission LUTs by the PSF.

2.2.3 Spectroscopic database

The Spectroscopic Database used in IPF 6.01 (or GOPR 7) is not up-to-date. The model transmission spectra LUTs have been computed using an old version (v2000) of the HITRAN spectroscopic database + the radiative transfer model LBLRTM. HITRAN spectroscopic database has been several time updated (2004, 2008, 2012) since then. Making use of an up-to-date version of HITRAN may have an impact on the quality of the GOMOS O\textsubscript{2} and/or H\textsubscript{2}O profiles retrieval.

Moreover, the LUTs used by IPF/GOPR do not take into account the Collision Induced Absorption (CIA) effect. CIA is spectral feature resulting of inelastic collisions of molecules in gas. As shown in Figure 8 for SPB1, it has a significant impact only at low altitude (below 15 km). We will add the possibility to take into account the CIA in the ALGOM O\textsubscript{2} processing chain. CIA also exists in SPB2 (925-955 nm) but has a much smaller absorption cross section than in SPB1. Considering its impact on H\textsubscript{2}O retrieval is negligible, we won’t take it into account in the ALGOM H\textsubscript{2}O processing chain.

A document has been written by S.Ferron [RD-3] which describes how the new model transmission spectra LUTs have been computed in the frame of the ALGOM project. This document also covers the computation of the CIA LUTs.
Figure 8: From top to bottom: model transmission in SPB1 spectral range (O\textsubscript{2} retrieval) for tangent point altitudes of 15, 10, 5 and 1 km without (black) or with CIA (red). Resolving power: 6000, US76. The influence of CIA is almost null at 15 km and above. Source: [RD-3]
3 New Algorithm

3.1 Description

A new algorithm has been developed for the spectral inversion in the frame of the ALGOM project. It is based on an inversion using a Levenberg Marquardt (LM) for fitting a model transmission spectrum to a measured transmission spectrum. This algorithm is very similar to the current Spectrometer A algorithm which uses also a Levenberg Marquardt. The new algorithm is basically the same for O₂ and H₂O retrieval. However some parameterization, for example fixing/letting-free some input parameters of the LM, may differ between H₂O and O₂ retrieval.

3.1.1 Inversion scheme

The inversion scheme uses a Levenberg Marquardt which retrieves the following parameters:

- **The logarithm of the slant path density \(N_{sl}\):** this is the main retrieved parameter which provides the slant path density profile after the fit convergence of each measurement.

- **The wavelength shift \(w_{\text{shift}}\):** this is an adjustment (shift) from the nominal wavelength assignment. The nominal wavelength assignment is provided in the GOM_EXT product but it is not accurate enough (see Figure 3 and Figure 4). It needs some adjustment which varies with star ID, season, etc. ...

- **The global PSF width \(\sigma_{\text{PSF}}\):** the global PSF (which includes the static PSF + dynamic PSF) sigma is retrieved. The model for PSF is a Gaussian centered on zero (no wavelength shift because wavelength shift is taken into account in a specific parameter). It is important to keep in mind that in most results presented in this document, we use HITRAN model transmission LUTs that have been already smoothed by applying a small PSF to the infinite resolution LUTs and then applying a sub-sampling. This means that in our results, the retrieved global PSF will not be the ‘true’ global PSF but the quadratic sum of the ‘true’ global PSF + this small PSF.

- **2 coefficients** for adjusting the baseline (continuum) of the transmission spectrum (one for the height \(\text{coef}_{\text{slant}}\), one for the slope \(\text{coef}_{\text{slope}}\)): these two coefficients allow to take into account effects such as Rayleigh or aerosol scattering which tend to offset the baseline of the transmission spectrum to values lower than 1.

The inversion by LM is performed for each measurement independently.

- **Measured transmission spectrum:** it is extracted from the GOM_EXT product: ‘transmission corrected from scintillation and dilution effects’;

- **Error bars of the measured transmission spectrum:** they are computed as the root square of the variance extracted from the GOM_EXT product: ‘covariance function of the transmission after scintillation and dilution effect correction’;
- **Flags**: the flags at sample level of the GOM_EXT products: ‘Flags for the transmission model’ are used. Flagged pixels are discarded of the fitting process.

**Spectral domain**

For O₂ retrieval, all pixels of SPB1 are used. We have also tested the possibility to add pixels of SPA2 around 690nm, covering the O₂ absorption band B. For H₂O retrieval, all pixels of SPB2 are used.

**Altitude domain**

O₂ retrieval is performed only for altitudes <85 km. H₂O retrieval is performed only for altitudes <50 km.

**Initialization and range of the parameters**

The parameter initialization value and allowed range are presented in Table 4. These are the default values used at the beginning of the projects. For specific tests, they may differ from the default values. If so, it shall be specified in the document where needed.

*Table 4: default configuration for the parameters of the levenberg-marquardt. The Allowed range is the range of tested values constrained by the Levenberg Marquardt. It never uses values which are outside of this range.*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initialization</th>
<th>Allowed range</th>
<th>Fixed/free</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N_{sl})</td>
<td>Climatological model (N_{sl_clim})</td>
<td>&gt; (N_{sl_clim}(100\text{km})), &lt; (N_{sl_clim}(0\text{km}))</td>
<td>Free</td>
</tr>
<tr>
<td>(w_{shift})</td>
<td>0</td>
<td>/</td>
<td>Free</td>
</tr>
<tr>
<td>(\sigma_{PSF})</td>
<td>O₂ (SPB1):0.025nm (SPA2):0.2nm H₂O:0.032nm</td>
<td>&gt;0.001 nm</td>
<td>Free</td>
</tr>
<tr>
<td>(\text{coef}_{slant})</td>
<td>0</td>
<td>/</td>
<td>Free</td>
</tr>
<tr>
<td>(\text{coef}_{slope})</td>
<td>0</td>
<td>/</td>
<td>Free</td>
</tr>
</tbody>
</table>

**3.1.2 Forward model**

The forward model takes as input the 5 parameters described in section 3.1.1 (reminded in red below). It performs these successive steps each time it is called by the Levenberg Marquardt:

1. Load the model transmission LUT for the climatological condition (season, latitude) of the processed occultation among 5 climatological models: tropical, midlat_summer, midlat_winter, subarc_summer, subarc_winter. The transmission LUT is provided for 100 altitudes (\(z = 1\text{km} \rightarrow 100 \text{km}, 1 \text{ km step}\) with one value of logarithm of density \(N_{sl_clim}(z)\) associated to each altitude.
2. Interpolation on logarithm of column densities: The transmission spectrum is interpolated linearly at the current (input of forward model) logarithm of density $N_{sl}$ between the two surrounding logarithm densities of the model transmission LUT $N_{sl,\text{clim}}(z)$ and $N_{sl,\text{clim}}(z+1)$.

3. Create the spectral PSF: A Gaussian function centered at 0 with a sigma equal to the $\sigma_{\text{PSF}}$.

4. Convolution of the transmission spectrum by the spectral PSF: two methods have been implemented for the convolution: the IDL function ‘convol’ or convolution by FFT (quicker for high resolution LUTs).

5. Transmission baseline correction: the baseline is adjusted in order to correct from Rayleigh and aerosol effects by using the 2 baseline parameters ($\text{coef}_{\text{slant}} + \text{coef}_{\text{slope}}$).

6. Integration on spectels.

7. Wavelength interpolation: the model transmission spectrum is resampled on the GOMOS wavelength grid (nominal wavelength assignment + $w/\text{shift}$) in order to be compared with the measured transmission spectrum.

### 3.1.3 Limitations

In IPF/GOPR the measured transmission spectra used in the SPB1 and SPB2 spectral inversion are corrected from O$_3$ absorption, Rayleigh and aerosol scattering. This is done before the fit presented in Figure 2. Note that the measured transmissions stored in the GOM_EXT product are not corrected from these effects. Indeed, this correction is performed by IPF/GOPR but only non-corrected spectra are saved in GOM_EXT. In ALGOM we do not apply this correction: we use the GOM_EXT measured transmission (which, as said above, are not corrected) in the LM. For Rayleigh and aerosol it is indirectly taken into account thanks to the baseline adjustment with $\text{coef}_{\text{slant}}$ and $\text{coef}_{\text{slope}}$ (note that this baseline adjustment also exists in IPF/GOPR and thus gives a supplement of correction if the Rayleigh and aerosol correction is not perfect). For O$_3$, which absorbs in SPB, nothing is done in ALGOM. We have investigated the impact of non taking into account the correction of O$_3$ in SPB. For this, we processed a dataset of 390 occultations of cold stars (ah_h2o dataset) in full dark illumination condition with GOPR with and without activating the O$_3$ correction. We computed the relative difference between the two versions for the O$_2$ and H$_2$O slant path densities profiles. Results are presented in the two figures below.
Legend of the figure above: Relative difference of O$_2$ tangent line density profiles between:

From the two figures above we can say that:
- The impact of non-taking into account O$_3$ correction in SPB2 is negligible on the H$_2$O retrieval: less than 0.5% which is very small compared to the desired accuracy for H$_2$O retrieval (of the order of 10%).
There is no impact of not taking into account the O₃ correction in SPB1 on the O₂ retrieval above 60 km. Below 60 km, there is a small impact that reaches -2% around 35 km. This is not negligible compared to the accuracy (random error) that Gomos is able to reach on the O₂ retrieval (at best 0.7% near 25 km) and compared to the accuracy of ecmwf data (about 2%). However, conclusions of this ALGOM study (see sections 5.3 and 7) have shown that GOMOS ALGOM O₂ retrieval suffers from unexplained strong biases compared to ecmwf, by more that 10%. Thus, the conclusions of our study would have stayed the same if we had added in ALGOM the O₃ correction in SPB. One shall keep in mind that if we understand in the future the strong bias vs. ecmwf (and can correct for it) then it will be important to add the SPB O₃ correction in the ALGOM processor.
4 ALGOM Dataset Generation

4.1 Format and content of the products

The ALGOM output products are in netCDF4/HDF5 format. There is one NetCDF file per occultation. The O\textsubscript{2} product and H\textsubscript{2}O product are in two separate files.

The content of the ALGOM O\textsubscript{2} and H\textsubscript{2}O output products have been harmonized as much as possible with the content proposed by FMI and ESA for ALGOM User Friendly Flagged products. Variables are sorted in groups.

An ALGOM NetCDF product contains the following variables stored in several groups:

**Geolocation group:**
- time
- Latitude.
- Longitude
- Time\textsubscript{start}
- Time\textsubscript{end}
- Latitude\textsubscript{start}
- Latitude\textsubscript{end}
- Longitude\textsubscript{start}
- Longitude\textsubscript{end}
- Altitude
- Altitude\textsubscript{min}
- Altitude\textsubscript{parameters}
- Duration
- Obliquity

**Radiation group:**
- Sza\textsubscript{tangentpoint}
- Illumination\_flag
- Sza\_satellite
- Saa\_flag

**Starttarget group:**
- Star\_id
- Star_temperature
- Star_magnitude
- Mean_flux_spectro_SPB1(2)_2003

**O2(H2O)_density group:**
- O2(H2O)_slant_density
- O2(H2O)_slant_density_std
- O2(H2O)_slant_density_confidence

**Retrieval_quality group:**
- Chi2

**External_atmospheric_model group:**
- Air_density_external

**Aprior_data group:**
- Transmission_model_lut_atmosphere
- Climatological_model

**Satellite_geolocation group:**
- Orbit_number
- Latitude_satellite
- Longitude_satellite
- Latitude_satellite_start
- Latitude_satellite_end
- Longitude_satellite_start
- Longitude_satellite_end

**metadata group:**
- title
- filename_netcdffile
- filenamesyntax
- GOM_EXT_source_file
- GOMOS_ALGOM_O2(H2O)_dataversion
- GOM_EXT_source_dataversion
- File_creation_date
4.2 Processing configuration

The three datasets were generated with the following ALGOM configuration parameters. These configuration parameters have been chosen because they seem to deliver profiles with the best achievable quality. One must refer to the Appendix for illustration/explanation on the impact of each configuration parameter on the quality of the profile.

*Table 5: Configuration parameters delivering the best profiles. These parameters have been used for generating the ALGOM sample datasets delivered to ESA.*

<table>
<thead>
<tr>
<th>ALGOM parameter</th>
<th>O2</th>
<th>H2O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral database for transmission model</td>
<td>HITRAN 2012</td>
<td>HITRAN 2012</td>
</tr>
<tr>
<td>Spectral pixels for LM fit</td>
<td>SPB1 only (O₂A absorption band)</td>
<td>SPB2</td>
</tr>
<tr>
<td>PSF width in LM</td>
<td>All stars except S004: fixed using pre-computed LUT</td>
<td>Fixed using pre-computed LUT*</td>
</tr>
<tr>
<td>Wavelength assignment in LM</td>
<td>free</td>
<td>free</td>
</tr>
<tr>
<td>A-priori atmosphere for model transmission LUTs</td>
<td>5 climatological models**</td>
<td>5 climatological models**</td>
</tr>
<tr>
<td>CIA correction</td>
<td>Included</td>
<td>/</td>
</tr>
<tr>
<td>Error bar correction factor</td>
<td>Not activated***</td>
<td>/</td>
</tr>
</tbody>
</table>

*S004 is discarded of H₂O retrieval because of too strong iPRNU structures (double star).
** It would be better to use the GOMOS external model but it takes much too much time to build the LUTs.
*** First results show that activating the correction factor helps to provide better error bars at low altitude but it needs more tests and validation before using it with confidence.
5 Project synthesis

This chapter summarizes the work performed during the project.

- We first summarize the differences in the algorithms between ALGOM and GOPR/IPF;
- Then, we summarize the differences in the quality of the retrieved profiles:
  - Improvement with ALGOM;
  - ALGOM unexpected results;
- Finally, we summarize the problems coming from Level-1B processing that we pointed out during this study.

5.1 Algorithms differences

The following elements are present in ALGOM spectral inversion processing chain but not in GOPR/IPF:

- Using a Levenberg Marquardt (LM) for fitting measurement with a model.
- Possibility to retrieve wavelength (wl) assignment as a free parameter of the Levenberg Marquardt. (Allows fitting seasonal variations of wl which is not the case in GOPR/IPF).
- Possibility to retrieve spectral PSF width as a free parameter of the Levenberg Marquardt (the PSF stored in ADF and used by GOPR/IPF is over-estimated).
- Use of a new spectroscopic database: HITRAN 2012 for building model transmission LUTs.
- Use of a model transmission LUT with a better spectral resolution than the ADF LUTs that GOPR/IPF use.
- Possibility to take into account Collision Induced Absorption (CIA).
- Possibility to build O$_2$ transmission LUTs from GOMOS external model profiles (but CPU intensive) instead of climatological model.
- Introduction of a transmission error bars correction factor for more realistic retrieved O$_2$ density error bars (especially at low altitude).
- Possibility to use the SPA2 O$_2$ absorption band B in addition to the SPB1 O$_2$ band A.

The following elements are present in GOPR/IPF but were not implemented in ALGOM:

- Measured transmission is corrected from O$_3$ absorption before the fit (O$_3$ is a weak absorber in SPB1, so this should not be significant).
- Measured transmission is corrected from Rayleigh and Aerosol before the fit (should be compensated in ALGOM by the fitting of the continuum in LM).
5.2 Improvement of the quality of the profiles

The improvements in the quality of ALGOM profiles compared to GOPR/IPF are:

- GOPR/IPF profiles stop at low altitude (when one spectral value of Tmeas <0 encountered). Not in ALGOM. Illustrated in Figure 9 (red circle).

- GOPR/IPF has many bad profiles at low altitude. All these bad profiles have disappeared in ALGOM. Illustrated in Figure 9 (green circle) and Figure 26.

- GOPR/IPF has a strange threshold for minimum value. Not in ALGOM. Illustrated in Figure 9 (black circle).

- O2 retrieval is of very poor quality in GOPR/IPF for faint stars. ALGOM is much better (even if still poor compared to bright stars retrieval). Illustrated in Figure 10.

- O2 and H2O dispersion is reduced with ALGOM. Illustrated in Figure 11 and Figure 12.

- Star S0004 (double star) is better handled by ALGOM thanks to the possibility to let the PSF width be a free parameter in LM. O2 profiles are improved with ALGOM. Illustrated in Figure 13.

- Choosing GOMOS external model as apriori for building (model transmission LUTs), instead of the 5 climatological models improves the quality of the O2 profiles above 35 km. Illustrated in Figure 14.

- Introducing a correction factor for the error bars of the measured transmission spectrum leads to more realistic retrieved O2 density error bars at low altitudes. Illustrated in Figure 15.
Figure 10: (Right and left plots) O₂ slant path density profiles for IPF/GOPR (blue) and ALGOM (red) for all occultations of faint star S0100. We see that GOPR/IPF provides unrealistic median (yellow thick line left plot) and percentile (yellow dotted line right plot) while for ALGOM (black lines) they are more realistic.

Figure 11: O₂ slant path density profiles divided by air ecmwf for IPF/GOPR (right plot) and ALGOM (left plot) for all occultations of faint star S0002. We see that the dispersion is smaller for ALGOM than for GOPR/IPF by about 25-30%. The strong bias at low altitudes has disappeared.
Figure 12: H$_2$O slant path density profile for all occultations of star S0002 for ALGOM (red) and GOPR/IPF (blue). We can see on the zoomed box that the dispersion is smaller in ALGOM than in GOPR/IPF.

Figure 13: Upper plot: ALGOM retrieved PSF width as a function of orbit number for all occultations of S0004. We clearly see periodical fluctuations due to the fact that S0004 is a double star, double image which total PSF changes with the variable inclination of the line joining the two stars with the horizontal (spectral direction). These fluctuations are not taken into account in GOPR/IPF. Lower plot: O$_2$ slant path density profiles divided by air ecmwf for IPF/GOPR (right plot) and ALGOM (left plot) for all occultations of faint star S0004. We see that the dispersion is smaller for ALGOM than for GOPR/IPF by about 40%.
5.3 ALGOM misunderstood/unexplained results

We have seen in section 5.2 several improvements of ALGOM compared to GOPR/IPF. There is however one unexpected ALGOM result: when performing the ratio of retrieved O$_2$ slant path density profiles with external model slant path air density profile, we find that for most stars GOPR/IPF median is generally closer to the expected value of 0.209 than ALGOM is (see Figure 16, Figure 19). Indeed, O$_2$ is a well mixed gas in the atmosphere and the mixing ratio is expected to be 20.9% at all altitudes. It is however interesting to notice that for star S0013, ALGOM is closer to the expected value than GOPR/IPF is (see Figure 21).
Note that the differences between GOPR/IPF and ALGOM results are understood and attributed mainly to the use of a different PSF width and wavelength assignment as explained in section 8.1.9. However, what we do not understand is why ALGOM is not closer to the expected 0.209 value compared to GOPR/IPF.

FOne should remind that the results presented in Figure 16 to Figure 21 have been obtained by using the model transmission LUTs computed from climatological models. According to the results presented in Figure 14, it is possible that using the model transmission LUTs computed from GOMOS external profile would reduce the desagreement between ALGOM O$_2$ retrieval and the external model especially in the altitude range [35-70 km].

The original goal of O$_2$ measurements was to get a vertical density profile of air and to measure the temperature from it, applying hydrostatic law to the vertical density profile. This was supposed to be done without any external information. But we found that even when using the external information of the vertical profile temperature, we have still some significant discrepancies, with the density of ECMWF which is in hydrostatic equilibrium.

We use the A band of O$_2$ at 760 nm, and we note that this band, used in GHG gases retrieval from space in nadir viewing is also not well fitting the meteo data. It means that there is a fundamental fact, not well understood both in occultation and in nadir viewing which remains not well understood. Further studies of the GOMOS ensemble could perhaps help to understand fully the problem, also of nadir viewing.

At present, the GOMOS O$_2$ data cannot be considered as validated, since the required accuracy is very high to be useful (<1-2%).

Figure 16: O$_2$ slant path density profiles divided by air ecmwf for IPF/GOPR (right plot) and ALGOM (left plot) for all occultations of faint star S0001. Green circled: We can see that above 35 km, the median is closer to the expected value of 0.209 for GOPR/IPF than for ALGOM.
Figure 17: Same as Figure 16 for star S0002. As for S0001, we can see that above 35 km, the median is closer to the expected value of 0.209 for GOPR/IPF than for ALGOM.

Figure 18: Same as Figure 16 for star S0016. As for S0001, we can see that above 35 km, the median is closer to the expected value of 0.209 for GOPR/IPF than for ALGOM.

Figure 19: Same as Figure 16 for star S0063. As for S0001, we can see that above 35 km, the median is closer to the expected value of 0.209 for GOPR/IPF than for ALGOM.
5.4 Level-1B problems

During this project we could point out the following limitations on the quality of the retrieved O$_2$ and/or H$_2$O profiles, coming from the Level-1B processing chain:

- Undetected cosmics can give outlier O$_2$/H$_2$O profiles. This is particularly the case in the SAA region where the number of cosmic hits is very high.
  - Solution/recommendation: Cosmic hit detection algorithm cannot handle spectra with too many cosmics hits. Just be careful when using Level-2 data observed in the SAA region.
- The bad dark charge correction, due to the fact that the DSA is too distant from the processed occultation tends to increase the dispersion of the O₂ profiles at low altitude (<20 km).
  - Solution/recommendation: Improve the quality of the dark charge correction: take the closest DSA + other method (DC in UV computed from last altitudes of the occultation, see dedicated Technical Note by L.Blanot [RD-5] written also in the frame of the ALGOM CCN project)

- The bad dark charge correction, due to the fact that the DSA is too distant from the processed occultation delivers under-estimated error bars on O₂ density profiles at low altitude (<20 km).
  - Solution/recommendation: Same as above + apply the error bar correction factor presented in this TN.

- Bad iPRNU correction can lead to bad H₂O profiles at 30-50 km.
  - Solution/recommendation: Improve the iPRNU correction by re-computing better star nominal position LUTs.

- Loss of the star tracking at low altitude for pcd_illum=4 (twilight + stray-light) occultations lead to wrong profiles for altitudes where the star is lost.
  - Solution/recommendation: Detect when the star tracking is lost (see [RD-4]; add a flag in the Level-1B product.

- Bad PRNU at band level correction associated with strong twilight/stray-light conditions (pcd_illum=4) leads to very bad estimated central background correction in SPB2 and thus completely unrealistic H₂O transmission spectrum. Note that the impact on the H₂O profile is not clear (needs more investigations).
  - This is a complicated point because it implies 2 corrections in the Level-1B processing: 1. The flat field at band level correction. 2. The ecb estimation and correction. Therefore, we do not consider this point as a priority in case of further work on GOMOS data.
6 Further improvements

The following points are proposed in order to improve further the quality of the O₂ and/or H₂O profiles:

**O₂ and H₂O profiles:**
- Flag the data in case of false star tracking (pcd_illum=4). (Level-1B). A work has already been started [RD-4] which consists in searching for a sudden break in the sfa angle elevation profile, which is a way to detect the loss of the star tracking.
- Find a way to decrease the CPU time for building model transmission LUTs from GOMOS external profile instead of using the ones built from the climatological models.

**H₂O retrieval:**
- Improve the iPRNU correction (Level-1B): for this, we need to recompute the star nominal position for the eight H₂O stars with a better method than the one that was used for building the current nominal position. The method has already been tested by L.BLANOT.
- Filter the residual iPRNU structures: for this we can apply a low frequency filtering by FFT on both measured transmission spectrum and model transmission spectrum.
7 General Conclusions

During this study, we developed a new method for the spectral inversion of O₂ and H₂O GOMOS profiles. The main change of the new method is to use a Levenberg Marquardt Fitting algorithm which is not the case in the official GOPR/IPF method. The advantage of using a Levenberg Marquardt is the possibility to set the spectral wavelength assignment and the spectral PSF width as retrieved parameters of the fit. Doing so we could:

- Take into account the seasonal variation of the wavelength assignment.
- Find that the spectrometer B1 PSF width used by GOPR/IPF is overestimated and thus use the correct one.

The O₂ and H₂O profiles obtained with the ALGOM breadboard are clearly improved by several aspects compared to the official IPF-v6 O₂ and H₂O profiles like for example an extended low altitude coverage, a reduced dispersion and a decrease of the number of outliers.

This study also pointed out that the retrieved O₂ and H₂O profiles are very sensitive to the a-priori temperature of the atmosphere chosen for building the model transmission LUTs. Indeed, it has been shown that the quality of the profiles is improved if we build the model transmission LUTs from gomos external model atmosphere (ecmwf+miss) instead of using the 5 climatological models.

However, despite the numerous improvements brought by ALGOM on the quality of the profiles, we still do not understand why the ratio between O₂ retrieved profiles and ecmwf air density profiles is generally closer to the expected value (0.209 for a well mixed O₂) for GOPR/IPF than for ALGOM. It might have to do with spectral data base inaccuracies. Systematic spectral differences should be investigated. On figure 1 there is a notable discrepancy data/model at 759.3 nm; the model for this figure was an older version of HITRAN. This discrepancy does not exist with the HITRAN 2012, but there might be others.

At present, the ALGOM O₂ data cannot be considered as validated, since the required accuracy is very high to be useful (<1-2%). Further studies of the GOMOS ensemble could perhaps help to understand fully the problem.
8 Appendix - Detailed results

This Appendix provides a detailed presentation of all tests (with associated results) that were performed during the ALGOM O\textsubscript{2} H\textsubscript{2}O project. This Appendix is divided in two parts: the first part concerns only O\textsubscript{2} retrieval and the second part only H\textsubscript{2}O retrieval.

8.1 O\textsubscript{2}

8.1.1 First results – comparison with GOPR

8.1.1.1 Fit and residual

We apply the ALGOM processing on a single occultation of Sirius (orbit= R09584, star_id= S0001, start time of the occultation = 30 Dec 2003 15h11mn58s). Figure 22, Figure 23 and Figure 25 show for three different tangent point altitudes (respectively 81 km, 25 km, and 6 km) the measurements spectra, the fitted model, and the residuals of the fit.

![Figure 22](image)

**Figure 22:** Upper plot: black: measured transmission spectrum at 81.27 km in SPB1 extracted from the GOM\_EXT product. Red: fitted model transmission spectrum by ALGOM processing. Blue: fitted model transmission spectrum by IFP/GOPR processing (extracted from the GOM\_EXT product). Lower plot: red and blue: residuals for respectively ALGOM and IFP/GOPR processing. We see that at 81 km the measured spectrum is dominated by the noise, even for a bright star like Sirius. The gap in the residual spectrum between 765 nm and 766 nm corresponds to pixels which have been flagged. These pixels are not used by the Levenberg Marquardt.
Figure 23: Same as Figure 22 for a measurement at 25.25 km of tangent altitude. At this altitude, the O₂ spectral lines are very clear. Note that for the pixels with O₂ lines, the residual has higher amplitude for IPF/GOPR than for ALGOM. This is a consequence of the overestimation of the static spectral PSF in IPF/GOPR (see Figure 24).
Figure 24: Same as Figure 23 but zoomed in the spectral dimension between 762 nm and 765 nm. We see that the IPF/GOPR modelled transmission lines are smoothed too much. Indeed, the PSF for IPF/GOPR is imposed and is over-estimated as explained in section 2.2.2 whereas ALGOM PSF width is a free parameter of the Levenberg-Marquardt. The chi² is considerably reduced with ALGOM from 6.01 to 2.31.
8.1.1.2 Retrieved profiles

Our first tests were performed on a dataset made of 190 occultations of Sirius (S0001) covering the whole mission (2002 -> 2012). Sirius is the brightest star in visible and the second brightest star in SPB1 (S0014 is the brightest). Moreover, we know from past-studies (see Figure 4) that Sirius occultations show obvious seasonal variations in the wavelength assignment. For these reasons, Sirius is a good star for testing the performances of the new algorithm. All parameters are free parameters of the Levenberg Marquardt. The model transmission LUTs are the one of the ADF (GOM_CRS product), used by IPF/GOPR and they do not include the CIA correction.

In the following figures, we compare results of the ALGOM O₂ retrieval (red) with the results of IPF/GOPR (blue) O₂ retrieval:

- Retrieved Slant path density profiles in Figure 26 and Figure 27 and comparison with ecmwf in Figure 28.
- Chi2 profile in Figure 29.
- PSF sigma profile in Figure 30.
- Wavelength assignment profile in Figure 31.
- Coefficients of the baseline profile in Figure 32.

The main conclusions of the analysis of this dataset (thus valid for bright stars) are:
+ $O_2$ profile dispersion above 40 km seems a little bit smaller for ALGOM.
+ IPF/GOPR fit systematically fails below about 13 km revealing a problem/bug in the processor.
+ PSF sigma can be retrieved by the LM with good precision in the range [10-70 km]
+ Wavelength assignment can be retrieved with good precision in the same range [10-70 km]
+ IPF/GOPR PSF width is over-estimated.
+ There is a negative bias of up to -20% above 40 km between ALGOM and IPF/GOPR.

- ALGOM $O_2$ profiles are less in agreement with ECMWF (expected mixing ratio of 0.209) than IPF/GOPR $O_2$ profiles.

More details are presented in the legend of the figures below.
Slant Path density

Figure 26: Upper plots: log of O$_2$ slant path density profiles for S001 retrieved by ALGOM (red) and by GOPR (blue). The left and right plots are the same but on the left plot IPF/GOPR points are plotted over ALGOM points while on the right plot ALGOM points are plotted over IPF/GOPR points. Note also that the left plot show the medians while the right plot show the percentiles (16%, 84%). Lower plot: relative difference of the O$_2$ slant path density profiles between ALGOM and IPF/GOPR. The relative difference profile shows a negative bias up to -25% at 40 km. An attempt to explain this bias is provided in section 8.1.9. We see that the GOPR density profiles systematically fail below about 13 km (vertical stripes) due to a problem in the processor (computation of the logarithm of negative transmission values). Those points explain the behaviour of the relative difference plot below 13km (huge positive bias). Those points are flagged by IPF/GOPR.
Figure 27: Error bars (absolute) on the logarithm of the slant path density profiles. There is a bug in the version 6 of the IPF level2 products: the error bars on slant path densities have been wrongly coded in the product and it is impossible to recompute them for SPB. As a consequence, this plot has been done by using products generated with the version 5 of IPF instead of version 6. We see that between 40 and 70 km, the error bars are smaller by about 20% for ALGOM.

Figure 28: The slant path density profiles of ALGOM (left) and IPF/GOPR (right) are compared to the ECMWF air density profiles. The expected value is 0.20948, the mixing ratio of O₂. It is not clear if ALGOM is closer to this value than IPF/GOPR. Indeed, above 40 km, it seems that there is a negative bias for ALGOM. We note also that the dispersion above 40 km seems lower for ALGOM than for IPF/GOPR. The divergence for IPF/GOPR below 13 km is due to the fit failure (see Figure 25).
Figure 29: Chi2 profile for ALGOM (red) and IPF/GOPR (blue). The chi2 is always smaller for ALGOM (and stays >1). This is mostly due to a better estimate of the PSF by ALGOM, leading to smaller residuals as shown in Figure 23 and Figure 24. Another contribution to this lower chi2 can be the better estimate of the wavelength assignment (see Figure 31). Below 13 km, IPF/GOPR fit fails (see Figure 25).
Figure 30: Lower plot. Left: profiles of the retrieved global spectral PSF sigma (median in black). Right: associated error bars. Upper plot: width of the global PSF at 24 km altitude as a function of the orbit number (2002 → 2012). We see on the upper plot that the retrieved PSF sigma at 24 km is about 0.024 nm at the beginning of the mission and slightly decreases to about 0.020 nm at the end of the mission. There is no seasonal variation of the PSF sigma. We see on the left lower plot that the PSF width starts to increase below 25 km due to increased scintillations. However below about 10 km, the quality of the sigma PSF retrieval decreases probably because of lower SNR and also because most lines become saturated (see Figure 25). On the same plot we can notice that the PSF width is very slightly lower at 25 km than above 40 km.
Wavelength assignment shift

Figure 31: Same as Figure 30 but for the wavelength assignment shift. On the upper plot, the nominal wavelength assignment (extracted from the GOM_EXT product) has been plotted in black. This nominal wavelength assignment is used by IPF/GOPR. ALGOM improves this wavelength assignment by adding a retrieved shift to this nominal value. We see on the upper plot that the retrieved ALGOM wavelength assignment has seasonal variations (kind of ‘V’ shape). The strong dispersion of the points on the wl shift profile plot (lower left) is due to these seasonal variations. If we look carefully at all individual occultations, it is possible to see that the wavelength shift slightly increases below 15 km.
Coefficients for the baseline

![Figure 32: Upper left: profile of the retrieved height baseline coefficient. Upper right: associated error bars. Lower left: profile of the retrieved slope baseline coefficient. Upper right: associated error bars.](image)

8.1.2 Dependence with Star brightness in SPB1

The O$_2$ slant path density profiles for datasets made of stars S0063 (7500 e- in SPB1), S0041 (4200 e- in SPB1), S0029 (1354 e- in SPB1), S0157 (476 e- in SPB1) are presented respectively in Figure 33, Figure 34, Figure 35 and Figure 36. S0001, presented in Figure 26, has 22300 e- in SPB1.

Main conclusions are:

- The fainter the star is in SPB1, the larger the dispersion of the O$_2$ profiles is (true for both ALGOM and IPF/GOPR).
- The fainter the star is the higher is the altitude where the IPF/GOPR fit starts to fail (data are analysed with decreasing altitudes). For example for S0029 all IPF/GOPR profiles below about 23 km fail (because of the log of negative value bug).
- For very faint star (S0157), GOPR profile is completely erratic whatever the altitude is (this has been verified on other very faint stars like S0106, not plotted here).
- For all the stars presented here, we find the same negative bias above 40 km between ALGOM and IPF/GOPR, already found for S0001 (see Figure 26).
Figure 33: Same as Figure 26 but for a dataset made of star S0063. We find the same negative bias between ALGOM and GOPR/IPF than for S0001 dataset. Below 10-15 km, ALGOM shows some outliers which are due to the low SNR at low altitude. Note the very small spread of values in the range 20-30 km of altitude, compared to star S0001 (figure 26). This is because star S0063 probes only intertropical latitudes while S0001 probes most latitudes, with a much larger natural latitude variations.
Figure 34: Same as Figure 26 but for a dataset made of star S0041. Note the very small spread of values in the range 20-30 km of altitude, compared to star S0001 (figure 26). Same comment as in Figure 33.
Figure 35: Same as Figure 26 but for a dataset made of star S0029. GOPR/IPF profiles fail systematically below about 23 km.
8.1.3 Fixing PSF width in LM

We computed PSF width LUTs for SPB1 as shown in Figure 37. As shown in Figure 30, the PSF width is slowly decreasing through the time of the mission. Therefore we computed 2 LUTs: one for the early mission and one for the late mission.

We then compared the O$_2$ retrieval between:

- Run1: fixing in the Levenberg Marquardt the PSF sigma with the LUTs presented in Figure 37.
- Run2: letting the PSF sigma free in the Levenberg Marquardt

Results are presented below.

**Conclusions**

There is no obvious improvement of the quality of the O$_2$ profiles when fixing the PSF width (using the pre-computed LUTs) in the Levenberg Marquardt.
Figure 37: Retrieved SPB1 PSFs sigma profiles for several stars (coloured lines) and mean of all stars (black line) for early mission (left) and late mission (right). The mean curves are stored and can be used when the ALGOM configuration is set on 'FIXED' PSF width in the Levenberg Marquardt. Interpolation in time and in altitude in these LUTs will be used.
Figure 38: Retrieved slant path density $O_2$ profiles divided by air ECMWF with PSF fixed (red) or PSF free (blue) in the Levenberg Marquardt for all occultations of star S0001 (upper), S0063 (lower). Left is red plotted over blue, right is blue plotted over left. Median is visible on left plot (thick line) and percentiles on right plot (dotted line). Colors for median and percentile is yellow for PSF fixed and black for PSF free. For the two stars, the difference between PSF free and PSF fixed is negligible except a small reduction of the dispersion for PSF fixed above 60 km.

8.1.4 Fixing wavelength assignment in LM

We have seen that there is a seasonal variation of the wavelength assignment which is different for all stars. Therefore it is not possible to compute one wl shift LUT which would be valid for all stars. However, what we can do is to retrieve the wavelength assignment at 25 km (this is the altitude at which the wavelength shift retrieval is the most accurate) and then apply a LUT in order to take into account the altitude variations. This LUT is represented in Figure 39.
Figure 39: Retrieved SPB1 wavelength shift profiles for several stars (colored lines) and mean of all stars (black line). The mean curve is stored and can be used when the ALGOM configuration is set on ‘FIXED’ WL in the Levenberg Marquardt. In this case, a retrieval of wavelength shift at 25 km with ‘FREE’ WL is performed, then the LUT is applied in order to find the wavelength shift at other altitudes.

We then compared the O₂ retrieval between:

- Run1: retrieve the wavelength assignment at 25 km + altitude dependant LUT (presented in Figure 37).
- Run2: letting the wavelength assignment free at all altitudes in the Levenberg Marquardt

Results are presented in Figure 40.
Figure 40: Retrieved slant path density O₂ profiles divided by air ecmwf with 25km+LUT retrieved Wavelength shift (red) or free wavelength shift (blue) in the Levenberg Marquardt for all occultations of star S0001 (upper), S0002 (middle) and S0029 (lower). Median is visible on left plot (thick line) and percentiles on right plot (dotted line). Colors for median and percentile is yellow for wl 25km+LUT and black for wl free. Left is red plotted over blue, right is blue plotted over left. The difference between blue and red points is negligible for the bright stars S0001 and S0002. There is a small difference for S0029 above 50 km but it is not obvious that the profiles are improved in one case compared to the other.
8.1.5 Dependence with HITRAN version

We tested the $O_2$ retrieval on the S0001 dataset with the model transmission LUTs presented below. All transmission LUTs have been created using the HITRAN cross section spectroscopic database + the radiative transfer model LBLRTM.

- **GOPR**: This LUT is the one used by GOPR/IPF. It is extracted from the GOM_CRS ADF product. The LUT is made of only 1251 wavelengths covering SPB1. We noticed that 1251 points is not enough to preserve a proper shape of the $O_2$ transmission lines. With this LUT, the narrower lines are defined on only one or two points. It results in a jaggy shape (see Figure 41). The spectroscopic database is HITRAN v2000 and the radiative transfer model is LBLRTM. The resolving power is about 35000@764.5nm.

- **HITRAN 2004**: The HITRAN 2004 transmission LUTs have been calculated in the frame of the ALGOM project with a resolving power of 35000@764.5nm (chosen the same as GOPR LUT) and is sampled on 8192 points. This number is sufficient to preserve the correct shape of the lines as shown in Figure 41. Moreover it is optimized for convolution by FFT because $8192 = 2^{13}$. The spectroscopic database is HITRAN 2004 and the radiative transfer model is LBLRTM v12.2.

- **HITRAN 2008**: same as HITRAN 2004 with the 2008 version of the HITRAN spectroscopic database.

- **HITRAN 2012**: same as HITRAN 2004 with the 2012 version of the HITRAN spectroscopic database.

Figure 42 shows the comparison of ALGOM retrieved $O_2$ slant path density profiles for GOPR LUT vs. HITRAN 2004 LUT. Figure 44 shows HITRAN 2004 LUT vs. HITRAN 2008 LUT. Figure 45 shows HITRAN 2004 LUT vs. HITRAN 2012 LUT.

We note that the differences in the $O_2$ retrieved profiles between HITRAN 2004, HITRAN 2008 and HITRAN 2012 are very small (<1%). However the differences between GOPR LUT and HITRAN 2004 are bigger (negative bias, up to -5 % at 40 km). For this comparison the Chi2 is smaller above 30 km when using the HITRAN 2004 LUT but is higher below 30 km (see Figure 43).
Figure 41: O$_2$ model transmission LUT at tangent height altitude of 20 km for different versions (2004, 2008 and 2012) of HITRAN smoothed to a resolving power of 35000@764nm, for HITRAN 2012 infinite resolution and for GOPR LUT (black). We can see that the sampling of the GOPR LUT is too low to preserve the shape of the lines (lines become jaggy). We see also that the wavelength of the lines is not the same between GOPR LUT and HITRAN LUTs. GOPR LUT lines wavelength is generally smaller than HITRAN by about 0.01nm but this is not always the case like for example the first line of this plot, near 764.03nm, which is located at the same wavelength for HITRAN and GOPR LUTs.
Figure 42: Comparison of O2 slant path profiles from S0001 dataset between using the GOPR transmission model LUTs vs. the HITRAN 2004 recomputed transmission model LUTs. There is a negative bias above 30 km reaching about -5% at 40 km.
Figure 43: Comparison of Chi2 profiles for O2 retrieval of Figure 42.

Figure 44: Same as Figure 42 but for HITRAN 2004 recomputed transmission LUTs vs. HITRAN 2008 recomputed transmission LUTs. The differences are small (<1%).
Figure 45: Same as Figure 42 but for HITRAN 2004 recomputed transmission LUTs vs. HITRAN 2012 recomputed transmission LUTs. The differences are small (<1%).

8.1.6 Dependence with CIA

We performed and compared two runs on the S0001 dataset: one without considering the CIA and one introducing the CIA correction.

Conclusion: Taking into account the CIA seems to improve the O₂ retrieved profiles below 15 km by reducing a low altitude bias with respect to ECMWF.

Remark: The CIA correction that we implemented is a zeroth-order approximation: the absorption features due to CIA are dependent of the current O₂ slant path density $N_{sl}$ and thus the correction should be implemented in the forward model and make use of the current $N_{sl}$ each time the forward model is called by the Levenberg Marquardt. In our simplified approach, we used the climatological O₂ slant path density $N_{sl,c}$ instead of $N_{sl}$. As a consequence, the CIA correction, which consists in adding the CIA absorption features to the HITRAN model transmission LUT, can be performed outside the forward model only once for the whole occultation. This zeroth-order correction is sufficient for a first investigation about the impact of the CIA by comparing the results to a non-CIA corrected run.
Figure 46: Comparison of two runs of O₂ retrieval on S0001 dataset: run 1 = without taking into account the CIA (blue in upper plots), run 2 = taking into account the CIA (red in upper plots). The lower plot shows the relative difference. As expected, the impact is only below 20 km where the ‘with CIA’ run retrieves less O₂ than the ‘without CIA’ run.
Figure 47: Chi2 profile and chi2 relative difference of the runs presented in Figure 46. When taking into account the CIA, the chi2 is slightly reduced below 20 km (up to -5% near 10-15 km).

Figure 48: Comparison of the PSF width (sigma) between the two runs presented in Figure 46. The lower right plot shows that when taking into account the CIA, the PSF width is slightly decreased (-2%) at altitudes (> 20 km) where we would not expect any influence of the CIA. The impact below 20 km is bigger: there is a negative bias up to -20% at 5 km.
8.1.7 Dependence with climatological model

The model transmission LUTs are provided for 5 climatological conditions (Tropical, Mid_Lat_summer, Mid_lat_winter, Subarctic_summer, subarctic_winter) + US standard 1976. In order to test the impact of the choice of the climatological model on the O₂ retrieved density profile, we performed the comparison between these two runs on S0001 dataset:

- nominal run (i.e. choosing the model according to the latitude and season)
- run with climatological model forced to US standard for all occultations of the dataset.

Results are presented in Figure 50. We see that the choice of the climatological model has a large impact on the profiles as two different climatological models can lead to O₂ slant path density differences of up to 50%. This is a major result as it points out the fact that the retrieval is strongly dependent to the temperature profile and that the current division of the LUTs into only five climatological models is not accurate enough to reflect the real temperature profile of the atmosphere.

Figure 49: Left: O₂ vs. ecmwf profile for a run without taking into account CIA. Right: same as left plot but for a run taking into account the CIA. We see that the break of the curve around 10-15km on the left plot has disappeared on the right plot. This result tends to prove that the CIA correction helps removing a bias in the O₂ profiles at low altitude.
Figure 50: Upper plots: O\textsubscript{2} slant path density profiles obtain either by taking the good (lat/season dependant) climatological model LUT (red) or by forcing the transmission model to US STANDARD (blue). Lower plot: relative difference profile. The colors correspond to the climatological conditions (see the legend in the upper right corner of the plot). We see that the retrieved O\textsubscript{2} profile is very sensitive to the choice of the climatological transmission LUT as different climatological profile can lead to O\textsubscript{2} slant path density differences of up to +50% around 40 km.

8.1.8 Build model transmission LUT from Gomos External profile

The results presented in section 8.1.7 show that the retrieval is very dependant of the temperature of the atmosphere and therefore that using model transmission LUTs built from a climatological model atmosphere is too approximative for a good O\textsubscript{2} retrieval.

We have implemented the possibility to use a model transmission LUT built from the GOMOS external model atmosphere (ECMWF + MSIS), thus much closer to the ‘real’ atmosphere state during the processed occultation than when using climatology. Therefore, in this new approach, each processed occultation has its own model transmission LUT.

Building one transmission LUT takes about 50 min CPU time (on one processor), thus we could test this new approach only on a limited number of occultations. We computed these new model transmission LUTs for a dataset of 191 occultations of Sirius (same dataset as in Figure 50). Results are presented in Figure 51.
Figure 51: O2 slant path density profiles divided by Gomos external model air density for a dataset made of 191 occultations of star S0001. Note that this star covers a wide range of latitude which is good for the current test. Blue: the model transmission LUTs come from climatological model. Red: the model transmission LUTs come from GOMOS External atmospheric model (ecmwf + msis). Median is visible on left plot (thick line) and percentiles on right plot (dotted line). Colors for median and percentile is yellow for GOMOS EXTERNAL and black for CLIMATOLOGY. Using the model transmission LUTs from GOMOS external model seems to improve the results because:
- the dispersion is reduced (check percentile on right plot) between 40 and 70 km.
- the negative bias in the median between 40 and 70 km is reduced.

Conclusions: Using a model transmission LUTs seem to improve the quality of the O2 profiles between 35 and 70 km. Note that this altitude range is consistent with the results presented in Figure 50.

8.1.9 Understanding the -20 % Negative Bias wrt IPF/GOPR

We have seen in Figure 26 that there is an ALGOM vs. IPF/GOPR negative bias on the O2 slant path densities above 20 km that reaches -20% around 40 km. This negative bias is also present on the retrieval of O2 profiles, using other stars, as shown in section 8.1.2.

We have investigated the impact of the PSF sigma on this bias. For this we performed several runs by forcing the PSF sigma (only above 25 km) to several values from 0.020 nm to 0.060 nm with a 0.005 nm step.

Results are presented in Figure 52 (O2 slant path density profiles) and in Figure 53 (chi2). We see that the ALGOM vs. IPF/GOPR negative bias is the strongest for the narrower PSF (0.020 nm) and that it progressively disappears when the PSF width increases. We also note that the chi2 increases and that it is the closest to the IPF/GOPR chi2 value for PSF sigma = 0.045 nm.

It is interesting to note that the same kind of behaviour occurs when fixing the Wavelength Assignment to several values [-0.03, -0.02, -0.01, 0, 0.01, 0.02, 0.0] nm, instead of the PSF sigma. Results are not presented in this document.
According to these results, we can say that the negative bias may be due to a difference in PSF sigma and/or wavelength assignment between IPF/GOPR and ALGOM. However, we still do not understand why the agreement with ECMWF air density is better for IPF/GOPR.

Figure 52: Relative difference of O2 retrieved profiles between ALGOM and IPF/GOPR, for S0001 dataset, for different values of the PSF sigma fixed in the Levenberg Marquardt (only above 20 km, below 20 km the PSF sigma is let free, which explains the break at 20km). We see that the negative bias present for small sigma PSF values tends to disappear when the sigma PSF increases.
Figure 53: Same as Figure 52 but for the Chi 2 profiles. Blue is IPF/GOPR, red is ALGOM. Above 20 km, where the PSF has been fixed, the chi2 is almost equivalent between IPF/GOPR and ALGOM for PSF sigma=0.045 nm which signifies that the global PSF sigma used by IPF/GOPR is probably close to 0.045 nm.

8.1.10 Improving the error bars

It has been shown in [RD-5] that when the dark charge is not properly corrected, then the transmission spectrum error bars are under-estimated at altitudes and on spectral pixels where the star signal level is lower than the residual dark charge level. This results in under-estimated slant path densities error bars.

We introduced a correction factor for an improved estimation of the transmission spectrum error bars. This correction factor is based on the standard deviation of the continuum of the O₂ transmission spectrum at wavelengths < 759 nm as shown in Figure 54.

The correction factor \( R = \frac{\text{standard deviation of transmission for pixels <759nm}}{\text{median of the error bars}} \).

The new error bars are then computed as the old ones multiplied by this correction factor.
Figure 54: Transmission spectrum at 37 km for a Sirius occultation. The correction factor is computed as the ratio between the standard deviation of the pixels of the continuum (red circle) by the median of the error bars of the same pixels.

Figure 55: Upper plots: retrieved O$_2$ slant path density profiles with associated error bars in green for the case without applying the correction factor (left) and the case applying the correction factor (right). Lower plots are a zoom for altitude < 20km. We see that with the correction factor applied, the error bars at low altitudes are increased and more realistic than without the correction factor.
Figure 56: Upper left: chi2 profiles for the case applying the correction factor (blue) and the case not applying the correction factor (red). Upper right: same as upper left with blue overplotted on red. Lower left: retrieved O2 density error bars. Lower right: same as lower left with blue overplotted on red. We can see that the chi2 is reduced and closer to 1 when the correction factor is applied. The error bars below 20 km are increased when correction factor is applied.

8.1.11 False star tracking

It has been shown in [RD-5] that in strong stray-light illumination conditions (pcd_illum=4, pcd_illum=1), for several occultations the star tracking fails at low altitude. In this case, something else than the star is tracked, but the measurements continue being performed. The altitude at which the star tracking is lost can be determined from the sfa elevation angle profile (see Figure 57). These occultations deliver bad O2 density profiles below the altitude where the star is lost as shown in Figure 58 for a single case and in Figure 59 for a larger sample of occultations.

The method proposed in [RD-5] for detecting a false star tracking should be implemented in the Level-1B processing chain for flagging these data.
Figure 57: Sfa azimuth profile for R32522 S0016 in pcd_iillum=4 illumination conditions. We see that at about 10 km, there is a break of the curve which is relevant of a false star tracking: the star tracker starts to follow something else than the star. A method is proposed in [RD-5] for automatically finding the presence of a break in the sfa elevation angle curve.

Figure 58: Left: O$_2$ slant path density profile for the case of Figure 57 (R32522 S0016) for ALGOM (red) and GOPR/IPF (blue). We see on the ALGOM profile that the retrieval becomes very bad below about 10 km which is due to the fact that the star is not tracked anymore. This is not reflected in the error bars (in black) which are still very small. GOPR/IPF profile fails below about 17 km (known problem already mentioned in this TN, not related with star tracking). Right: error bars of retrieved ALGOM O$_2$ density profiles.
8.1.12 Dark Charge correction

It has been shown in [RD-5] that when the DSA chosen for the Dark Charge (DC) correction is too far from the processed occultation, the retrieved O\(_3\) profiles can be of very poor quality. There is also an impact on the quality of the O\(_2\) profiles as shown on Figure 60 in which we can see that the dispersion of the O\(_2\) profiles at low altitude is higher when the DSA is far.
Figure 60: Upper left: ALGOM O₂ chi2 profile for dataset made of 579 products of star S0063 close in time (between R42000 and R44000). We see that there are two groups of points: first group (colored in red) with small chi2 value, second group (colored in blue) with much higher chi2 below 30 km. For the red points the DSA chosen for DC correction in the Level-1B processing is very close to the processed occultation (typically less than 5 minutes) while for the blue points it is much further (typically about 1h30, one orbit away). Upper right: slant path O₂ density profiles at altitudes <13 km for the same dataset. We see that the blue points (high chi2) are more dispersed than the red points. Lower right: errors bars of the slant path O₂ density profiles at altitudes <13 km. We see that the higher dispersion of the blue points in the O₂ profile is not reflected in the error bars which are roughly the same for blue and red points. Examples of transmission spectrum at low altitude for the two groups of points are provided in Figure 61.
Figure 61: Exemple of two measured transmission spectrum at about 8 km for a case when the dark charge correction is good (upper) and a case when the dark charge correction is bad (lower). In the upper plot, the DSA used for DC correction is very close in time from the occultation. In the lower plot the DSA used for the DC correction is far from the occultation. The upper plot corresponds to an occultation in the red group of Figure 60 while the lower plot corresponds to an occultation in the blue group.

We modified GOPR to version 80da in order that it can systematically take the closest DSA for the dark charge correction. We ran the same dataset as in Figure 60 with 80da in order to generate GOM_EXT products, then we processed the spectral inversion with ALGOM. Results are presented in Figure 62 for chi2 profile and in Figure 63 for the O2 slant path densities.

Conclusions: Choosing systematically the closest DSA improves the quality of the O2 retrieval by reducing the dispersion at low altitude. However, the improvement is moderate because the closest available DSA is often already too far in time from the processed occultation and thus does not deliver a very good DC correction.
Figure 62: Upper left: chi2 profile for ALGOM O2 spectral inversion for a dataset of 579 occultations of S063 between R42000 and R44000. The Level-1B processing comes from IPF-6 for the blue points and GOPR80da for the red points. Upper right: same as upper left with blue points plotted over red points. Lower: relative difference between blue and red points (reference is blue=IPF6). In the lower plot, below 30 km we see 3 groups of points: 1rst group is close to 0% which is relevant of points for which the same DSA was used in IPF6 and in GOPR80da. Therefore both IPF6 and GOPR80da could choose the closest available DSA. 2nd group is the one that goes around -40% below 20 km (with a high dispersion). For this group, the DSA used by GOPR80da is closer to the occultation than the DSA used by IPF6, however this closest available DSA is still at about 15-20 min from the occultation. The 3rd group is the one that goes at about -75% below 20 km. For this group of point we go from a very far DSA (about 1h30) used by IPF6 to a very close DSA (<5min) used by GOPR80da. The slant path densities profiles are presented in Figure 63.
Figure 63: $O_2$ slant path density profiles at altitude < 17 km for the dataset of 579 occultations of S063 between R42000 and R44000. The Level-1B processing comes from IPF-6 for the blue points and GOPR80da for the red points. Upper right: same as upper left with blue points plotted over red points, 5% 95% percentiles are in thick dotted line (black for IPF6, Yellow for GOPR80da). Median is visible on left plot (thick line) and percentiles on right plot (dotted line), Colors for median and percentile is yellow for 80da and black for IPF6. We see a reduction of the dispersion especially below 10 km for GOPR 80da compared to IPF6. This is due to the fact that for some occultations GOPR80da could find a closest DSA than IPF6.

8.1.13 Use SPA2 $O_2$B absorption band

The $O_2$ B absorption band is located at the end of SPA2 at 688-690 nm and is barely resolved spectrally, manifested as two absorption lines feature (the second one is incomplete). This absorption line is not used in IPF/GOPR spectral inversion for $O_2$ retrieval. We tested the usage of the $O_2$ B absorption line with ALGOM. First, we tried to retrieve $O_2$ with the $O_2$ B absorption band only, then we tried the retrieval using both $O_2$ B and $O_2$ A (SPB1 absorption band).

Conclusions: Using $O_2$ B (in SPA2) absorption band in addition to $O_2$ A (in SPB1) absorption band does not seem to obviously improve the quality of the retrieved $O_2$ profiles. However we notice a small reduction of the error bars on the retrieved $O_2$ profiles above 40 km when using both $O_2$A and $O_2$B absorption bands.
8.1.13.1 SPA2 $O_2 B$ only

Figure 64: Upper plot: black: measured transmission spectrum at 60.35 km in the last pixels (high wavelengths) of SPA2 extracted from the GOM_EXT product. Red: fitted model transmission spectrum by ALGOM processing. Blue: fitted SPA model transmission spectrum by IPF/GOPR processing (extracted from the GOM_EXT product). $O_2$ is not taken into account in this fit. Lower plot: red and blue: residuals for respectively ALGOM and IPF/GOPR processing. The $O_2 B$ absorption line can be seen at 687-688 nm but at this altitude the SNR is still very low, even for a bright star like Sirius.
Figure 65: Same as Figure 64 at 22.15 km. The O$_2$B absorption line is very clear seen between 686 and 690 nm.

Figure 66: Retrieved SPA2 PSF sigma for all occultation of Sirius from the inversion using the O$_2$B absorption band. We build a look up table (thick orange line) in order to be able fix the PSF width in LM.
Figure 67: Upper: retrieved O2 slant path profile for all occultations of S0002 using SPB1/O\textsubscript{2}A absorption band only (red) and SPA2/O\textsubscript{2}B absorption band only (blue). Lower: retrieved O2 slant path profile divided by air ecmwf. Median is visible on left plot (thick line) and percentiles on right plot (dotted line). Colors for median and percentile is yellow for SPB1/O\textsubscript{2}A and black for SPA2/O\textsubscript{2}B.
Figure 68: Same as Figure 67 for S0029. We see that the median of the ratio O2/air starts to diverge above 35 km.
8.1.13.2 SPA2 $O_2B + SPB1 O_2A$

Figure 69: Same as Figure 64 for combined $O_2A + O_2B$ retrieval at 25.35 km.
Figure 70: Upper plots: O$_2$ slant path density profiles obtain with O$_2$A (SPB1) absorption band only (red) and with O$_2$A (SPB1) + O$_2$B (SPA2) absorption bands (blue). Lower plot: relative difference profile. Dataset is about 200 occultations of star S0002. We do not see obvious differences between the two kinds of retrieval. The comparison with ECMWF is presented in Figure 71.
ALGOM: Improvements for GOMOS
O2 and H2O profiles retrieval
Technical Note

Figure 71: retrieved O2 slant path density profile divided by air ecmwf obtained with O2A (SPB1) absorption band only (red) and with O2A (SPB1) + O2B (SPA2) absorption bands (blue). Median is visible on left plot (thick line) and percentiles on right plot (dotted line), Colors for median and percentile is yellow for SPB1/O2A and black for O2A (SPB1) + O2B (SPA2). Dataset is about 200 occultations of star S0002. We do not see obvious differences between the two kinds of retrieval.

Figure 72: Error bars for the O2 profiles presented in Figure 70. We notice a small reduction of the error bars above 40 km when using both O2A and O2B absorption bands.

© 2016 ACRI-ST
Figure 73: Same as Figure 70 for star S0029. We do not see obvious differences between the two kinds of retrieval.
8.2 H2O

8.2.1 First results – comparison with GOPR

8.2.1.1 Fit and residual

We apply the ALGOM processing on a single occultation of Sirius (same star as Figure 22: orbit= R09584, star_id= S0001, start time of the occultation = 30 Dec 2003 15h11mn58s). Figure 75, Figure 76 and Figure 77 show for three different tangent point altitudes (respectively 49 km, 24 km, and 7 km) the measurements spectra, the fitted model, and the residuals of the fit.

From these figures we can say that:

- At 50 km the spectrum is dominated by the shot noise, we cannot see the H2O transmission lines.

- At 25 km, the spectrum is dominated by strong insufficiently corrected iPRNU structures. However it is surprising to see that the Levenberg Marquardt succeeds in fitting the measured spectrum by a model.

- Contrary to O2/SPB1 where the spectrum has the best SNR around 25 km, for H2O the best spectra are measured at the lowest altitudes of the occultation. Moreover, we see in Figure 77 that at 7 km H2O lines are still not saturated whereas most O2 lines are saturated at this altitude (see Figure 25).
Figure 75: Upper plot: black: measured transmission spectrum at tangent altitude = 49.12 km in SPB2 extracted from the GOM_EXT product. Red: fitted model transmission spectrum by ALGOM processing. Blue: fitted model transmission spectrum by IPF/GOPR processing (extracted from the GOM_EXT product). Lower plot: red and blue: residuals for respectively ALGOM and IPF/GOPR processing. We see that at 49 km the measured spectrum is dominated by the noise, even for a bright star like Sirius.
Figure 76: Same as Figure 75 for a measurement at 24.16 km of tangent altitude. At this altitude, the measured transmission spectrum is severely contaminated by residual iPRNU structures with amplitude higher than the H$_2$O lines. This shows that the iPRNU correction is insufficient at this altitude even though it still performs better than the nominal PRNU correction (used in IPF v5).
Figure 77: Same as Figure 75 for a measurement at 7.34 km of tangent altitude. At this altitude, the measured transmission spectrum shows some strong H$_2$O absorption lines. The residual is very close between ALGOM and GOPR/IPF.

### 8.2.1.2 Retrieved profiles

As for O$_2$ retrieval, our first tests were performed on a dataset made of 190 occultations of Sirius (S0001) covering the whole mission (2002 -> 2012). All parameters are free parameters of the Levenberg Marquardt. The model transmission LUTs are the one of the ADF (GOM_CRS product), used by IPF/GOPR.
**Slant Path density**

Figure 78: Upper plots: log H$_2$O slant path density profiles retrieved by ALGOM (red) and by GOPR (blue). The left and right plots are the same except on the left plot IPF/GOPR points are plotted over ALGOM points while on the right plot ALGOM points are plotted over IPF/GOPR points. Note also that the left plot shows the medians while the right plot shows the percentiles (16%, 84%). Lower plot: relative difference of the H$_2$O slant path density profiles between ALGOM and IPF/GOPR. The relative difference profile shows a negative bias up of about -15% above 15 km. We see that some GOPR density profiles fail below about 10 km (blue vertical stripes) due to a problem in the processor. This problem exists also for O$_2$ retrieval (see Figure 26). It seems that there are more outliers in ALGOM above 15 km (and especially above 40 km) than for IPF/GOPR. This is due to the fact that the PSF width is kept free for ALGOM (for this test) while it is fixed (to a rather correct estimation) for IPF/GOPR. Figure 97 presents the same plots for the ALGOM-delivered H$_2$O dataset adding a median plotted on the lower plot.
Figure 79: Error bars (absolute) on the logarithm of the slant path density profiles. There is a bug in the version 6 of the IPF level2 products: the error bars on slant path densities have been wrongly stored and it is impossible to recompute them for SPB. This is why only ALGOM error bars (in red) appear in this plot.

Chi2

Figure 80: Chi2 profile for ALGOM (red) and IPF/GOPR (blue). Above 10 km, ALGOM and IPF/GOPR show very close chi2. Below 10 km, some IPF/GOPR fit fails as explain above in this document.
**PSF width**

Figure 81: Lower plot. Left: profile of the retrieved global spectral PSF sigma of spectrometer SPB2. Right: associated error bars. Upper plot: sigma of the global PSF at 24 km altitude as a function of the orbit number (2002 -> 2012). The retrieval of the SPB2 PSF sigma is much less precise than for the one for SPB1 (see O2 retrieval, Figure 30). The decrease of the PSF width when the altitude increases is probably a bias introduced by the values for which the retrieved PSF width = 0.001 nm (minimum allowed PSF value by the LM). The higher is the altitude the more of these values there is because absorption features become smaller when altitude increases. These points should be eliminated because they probably correspond to unsuccessful fits.
Wavelength assignment shift

Figure 82: Same as Figure 81 but for the wavelength assignment shift. On the upper plot, the nominal wavelength assignment (extracted from the GOM_EXT product) has been plotted in black. This nominal wavelength assignment is used by IPF/GOPR. ALGOM improves the wavelength assignment by adding a possible shift to this nominal value. We see on the upper plot that the wavelength assignment has seasonal variations (same kind of ‘V’ shape than for SPB1 (see Figure 31)). The strong dispersion of the points on the wl shift profile plot (lower left) is due to these seasonal variations.
Coefficients for the baseline

Figure 83: Upper left: profile of the retrieved height baseline coefficient. Upper right: associated error bars. Lower left: profile of the retrieved slope baseline coefficient. Upper right: associated error bars.

8.2.2 Fixing PSF width in Levenberg Marquardt.

As for SPB1/O2 we built 2 LUTs for the PSF of SPB2. These LUTs are presented in Figure 84.

We then compared the H₂O retrieval between:

- Run1: fixing in the Levenberg Marquardt the PSF sigma with the LUTs presented in Figure 84.
- Run2: letting the PSF sigma free in the Levenberg Marquardt

Results are presented in Figure 85.

Conclusions: Contrary to the SPB1/O₂ retrieval the quality of the SPB2/H₂O retrieval is strongly improved when fixing the PSF width. This is due to the fact that when letting the PSF width free, it tends to fit strong iPRNU structures in the measured transmission spectrum instead of fitting H₂O absorption lines. This is especially the case above 15 km where the absorption lines are small compared to the iPRNU structures.
Figure 84: Same as Figure 37 for SPB2/H₂O. Above 35 km the SNR of the spectrum is too small giving over-estimated retrieved PSF sigma. Consequently we take a constant value above 35 km for our LUTs.
Figure 85: Retrieved slant path density H₂O profiles with fixed PSF width (red) or free PSF width (blue) in the Levenberg Marquardt for all occultations of star S0001 (upper), S0014 (middle) and S0063 (lower). Left is red plotted over blue, right is blue plotted over left. Letting the PSF free provides many wrong profiles above 15 km for the three stars. The cause is the iPRNU structures which are fitted by the PSF instead of the H₂O absorption lines. Fixing the PSF to the expected LUT value removes a lot of these bad profiles. We can see however that even when fixing the PSF there are still a few profiles that are biased like for example S0014 above 35 km (see Figure 96 for a dedicated figure on this feature). It is important to note that these results are not an improvement compared to GOPR/IPF which uses also a fixed PSF width.
8.2.3 Fixing wavelength assignment in LM

We apply the same test for H$_2$O retrieval as we applied in section 8.1.4 for O$_2$ retrieval. The wavelength altitude dependence LUTs is illustrated in Figure 86 and the comparison in the retrieved H$_2$O profiles are presented in Figure 87.

Conclusions: using the retrieved 25 km + LUT configuration seems to provide slightly less H$_2$O than using the free wavelength assignment at all altitudes configuration. However we cannot say which one is closer to the real concentration of H$_2$O (it would need some validation activity by comparing with other instrument). We can notice however that above 40 km the 25 km + LUT configuration gives a median which seems in a better continuity with the median below 40 km.

![Figure 86: Same as Figure 39 for SPB2/H2O.](image_url)
8.2.4 Using retrieved SPB1 wavelength assignment for SPB2 inversion

Figure 88 shows the correlation (scatter plots) between the retrieval of wavelength assignment sigma in SPB1 and the retrieval on exactly the same dataset of wavelength assignment in SPB2. The dataset used is all occultations of star S0014. Same plot for all occultations of star S0001 is shown in Figure 89.
Figure 88: Upper left: ALGOM retrieved wavelength assignment as a function of orbit number at 23.5 km for pixel 200 of SPB1 (blue) and pixel 200 of SPB2 (red) obtained respectively from ALGOM O$_2$ inversion and ALGOM H$_2$O inversion. The dataset used is all occultations of star S0014. Lower left: difference between SPB2 and SPB1 retrieved wavelength assignment. Right: scatter plot of SPB2 vs SPB1 wavelength assignment. We see a trend that starts at about orbit 13000 + some seasonal variations. The color refers to the orbit number (same color code as in lower left plot). We see that the scatter plot follows a linear model but that the intercept of this linear model varies with time. This variation of the intercept is the trend that we see in the lower left plot.

Figure 89: Same as Figure 88 for star S0001: results are very similar to S0014.
We can see that there is a strong correlation between \( \text{wl\_SPB2} \) and \( \text{wl\_SPB1} \). Therefore we built a model of \( \text{wl\_SPB2} = f(\text{wl\_SPB1}) \) as presented in Figure 90. This model is linear with an intercept which changes with time. The limitation of this model is that the seasonal variations are not taken into account.

We have seen in this document that the retrieval of the wavelength assignment is of better quality for SPB1/O\(_2\) than for SPB2/H\(_2\)O (because the O\(_2\) absorption lines are thinner and deeper than the H\(_2\)O absorption lines). Therefore, now that we know that \( \text{wl\_SPB1} \) and \( \text{wl\_SPB2} \) are strongly correlated, it could be interesting to use the SPB1/O\(_2\) output as a fixed input of SPB2/H\(_2\)O retrieval. Thus, we performed the H\(_2\)O retrieval by using wavelength assignment computed by this model (i.e. derived from the SPB1 O\(_2\) wavelength assignment). We then compared the retrieved profiles with the profiles obtained using a wavelength assignment retrieved directly from SPB2/H\(_2\)O. Results are presented in Figure 91.
Figure 91: Upper plots: H$_2$O slant path density profiles for all occultations of star S0063. Blue: wavelength assignment was fixed to a value computed using the SPB1 $\rightarrow$ SPB2 wl model, red: wavelength assignment is directly retrieved from the SPB2 H$_2$O inversion. We see that there is almost no difference between the two kinds of retrieval. Lower plot: relative difference profile between blue and red points.

Conclusions: We do not see any improvement in using SPB1/O$_2$ retrieved wavelength assignment for SPB2/H$_2$O inversion. We have also tried this method for fainter H$_2$O stars and do not see any significant improvement either.

8.2.5 Dependence with HITRAN version

As for O$_2$ retrieval, we test the H$_2$O retrieval with different version of the model transmission LUTs:

- GOPR LUT
- HITRAN 2004
- HITRAN 2008
- HITRAN 2012
Conclusions:

- Contrary to the $O_2$ retrieval, there is no obvious bias between GOPR LUT and HITRAN 2004 LUT. However, there is still a strong dispersion of the relative difference profile.
- Difference between HITRAN 2004 and HITRAN 2008 is small (<2%)
- HITRAN 2004 vs. HITRAN 2012 shows a negative bias of about -2.5% at all altitudes.

Figure 92: Upper plots: $H_2O$ slant path density profiles obtained with 2 different model transmission LUTs: GOPR LUT in blue, HITRAN 2004 in red. Lower plot: relative difference profile. Contrary to $O_2$ results, there is no negative bias. However the dispersion is still large.
Figure 93: Same as Figure 92 but comparing HITRAN 2004 and HITRAN 2008 model transmission LUTs. There is not a big impact.
8.2.6 Dependence with climatological model

The same test as for O\textsubscript{2} (see section 8.1.7) is performed for H\textsubscript{2}O in order to see the importance of the choice of the climatological model.

Conclusions: The H\textsubscript{2}O retrieval is also very dependant of the climatological model (leads to relative differences of up to -20\%). It means that the H\textsubscript{2}O retrieval is sensitive to the vertical actual temperature profile. But it would be perfectly acceptable, in the case of H\textsubscript{2}O, to use an external piece of information, the ECMWF pressure/temperature temperature profile of the time and place of occultation, which is in the product. The required accuracy for H\textsubscript{2}O is much less than for O\textsubscript{2}, and values with 10% error would already be a good achievement with respect to other H\textsubscript{2}O instruments.
Figure 95: Upper plots: H$_2$O slant path density profiles obtained either by taking the good (lat/season) climatological model LUT (red) or by forcing the transmission model to US STANDARD (blue). Lower plot: relative difference profile. The colors correspond to the Climatological conditions (see the legend in the upper right corner of the plot). As for O$_2$ retrieval, H$_2$O retrieval is also very sensitive to the choice of the climatological model. Here there is a bias of up to -20% at low altitude.

8.2.7 Problem with iPRNU correction

We know that the iPRNU correction is imperfect as shown in [RD-6], leading to residual fluctuations in the transmission spectrum. It seems that in the case where the correction is very bad, this has a strong impact on the quality of the O$_2$ retrieved profiles as shown in Figure 96.
8.2.8 Comparison ALGOM vs GOPR for the ALGOM-delivered H$_2$O dataset

Figure 97 shows the comparison of H$_2$O profiles between ALGOM and GOPR/IPF for the ALGOM-delivered H$_2$O dataset. It shows a negative bias of about -10% above 15 km for ALGOM. ALGOM processor has been configured with the parameters of Table 5.
Figure 97: Upper plots: log H$_2$O slant path density profiles retrieved by ALGOM (red) and by GOPR (blue). The dataset used is the ALGOM-delivered H$_2$O dataset, composed of the 8 brightest NIR stars (S004 excluded). ALGOM processor has been configured with the parameters of Table 5. The left and right plots are the same except on the left plot IPF/GOPR points are plotted over ALGOM points while on the right plot ALGOM points are plotted over IPF/GOPR points. Note also that the left plot shows the medians while the right plot shows the percentiles (16%, 84%). Lower plot: relative difference of the H$_2$O slant path density profiles between ALGOM and IPF/GOPR. The relative difference profile shows a negative bias with median of -10% above 15 km.

End of Document